

## Research Article

# Aero Engine Gas-Path Fault Diagnose Based on Multimodal Deep Neural Networks

Liang Zhao,<sup>1,2</sup> Chunyang Mo,<sup>1,2</sup> Tingting Sun,<sup>3</sup> and Wei Huang<sup>4</sup> 

<sup>1</sup>School of Software Technology, Dalian University of Technology, Dalian 116600, China

<sup>2</sup>Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian 116600, China

<sup>3</sup>Department of Natural Resources Information, Dalian Natural Resources Affairs Service Center, Dalian 116011, China

<sup>4</sup>First Affiliated Hospital of Dalian Medical University, Dalian 116000, China

Correspondence should be addressed to Wei Huang; [huangwei9898@163.com](mailto:huangwei9898@163.com)

Received 9 July 2020; Revised 5 August 2020; Accepted 21 September 2020; Published 6 October 2020

Academic Editor: Xiaojie Wang

Copyright © 2020 Liang Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aeroengine, served by gas turbine, is a highly sophisticated system. It is a hard task to analyze the location and cause of gas-path faults by computational-fluid-dynamics software or thermodynamic functions. Thus, artificial intelligence technologies rather than traditional thermodynamics methods are widely used to tackle this problem. Among them, methods based on neural networks, such as CNN and BPNN, cannot only obtain high classification accuracy but also favorably adapt to aeroengine data of various specifications. CNN has superior ability to extract and learn the attributes hiding in properties, whereas BPNN can keep eyesight on fitting the real distribution of original sample data. Inspired by them, this paper proposes a multimodal method that integrates the classification ability of these two excellent models, so that complementary information can be identified to improve the accuracy of diagnosis results. Experiments on several UCR time series datasets and aeroengine fault datasets show that the proposed model has more promising and robust performance compared to the typical and the state-of-the-art methods.

## 1. Introduction

Aeroengine is known as “the pearl on the crown of industry” because of the irreplaceable roles it plays in industry and the highly sophisticated internal structure it has. The fault diagnosis technology of aeroengine is vitally important to guarantee its performance and efficiency and reduce the maintenance cost, which can accurately diagnose and locate the fault of aeroengine and provide powerful support for engine maintenance. Therefore, to a certain extent, the efficiency decline and performance instability of engine caused by component fault may become predictable. On the other hand, the complexity of aeroengine can be expressed by the complex nonlinear relationship among thermodynamic parameters of its components. The fault and degradation of aeroengine components can reflect in thermodynamic parameters such as temperature and pressure. However, it is difficult to reverse this process, which means to deduce the causes and locations of faults by traditional computa-

tional fluid dynamics (CFD) software tools or by thermodynamics and fluid mechanics partial differential equation. Thus, most of the existing fault diagnosis technologies are realized by big data and artificial intelligence methods, which do not analyze why fault occurs but utilize the features or distribution of data to identify fault modes.

The existing aeroengine fault diagnosis methods can be roughly summarized as the following categories. First, techniques based on signal processing, such as the applications of Wavelet transform, Fourier transform, and Kalman filter. These methods extract features from the continuous signal waveform and present diagnosis results by analyzing them [1]. However, they are only applicable to continuously sampled data. Second, classification or clustering algorithms are based on similarity and distance. They treat each dimension of sample data as a coordinate of multidimensional vectors and find boundaries or curves in feature space to split samples into different groups. Data used by these methods should have strong numerical differentiation. An example is SVM,

which has been optimized into various versions [2, 3] from the basic form [4] or used in combination with other models [5, 6], in order to better cope with this problem. K-means algorithm [7] based on Euclidean distance to cluster states of the engine belongs to this category, as well. Third, diagnosis methods by combining expert experience and mathematical models. Expert experience is usually expressed as a decision tree or table which generates branches according to the threshold of sample attributes or features. In addition, expert experience is usually strongly subjective, so its correctness will directly affect diagnosis results. Fourth, diagnosis models based on neural networks, such as CNN, BPNN, DBN [8], PNN [9], and other deep learning models. The first three groups of methods mentioned above work merely by separating data or their features, without carefully fitting the functional or causal relationship between sample data and its categories, while methods based on neural networks can tackle it.

Therefore, in recent years, the deep learning models, such as multilayer autoencoders, CNN, LSTM, DBN, and BPNN, have attracted more attentions in the field of aeroengine fault detection and repair. Where CNN has the ability of learning features and is a powerful vehicle in many fields including problems related to sequence data. Thus, it was brought into the field of fault diagnosis, aiming to distinguish samples by its features. BPNN is a basic neural network that has the ability to express complex multivariate functional relationship. Hence, it is competent to most classification and regression tasks in various fields [10]. Due to its widespread applicability, it is now often used as the baseline in this field, or as the classifier in combination with other models.

In summary, CNN has the talent in extracting features from input data and then classifies data by them, while BPNN focuses on fitting original sample data and has a wide range of applications. However, most of the existing fault diagnosis mechanisms are designed only based on one neural network, which does not take advantage of the characteristics of both CNN and BPNN. Therefore, this paper proposes a multimodal deep neural network diagnosis method based on the feature perception ability of CNN and the fitting ability of BPNN. Specifically, the BPNN is employed to fit the sample data distribution, and CNN is utilized to explore the features of the samples, so as to obtain the multimodal decision information learned from different angles [11]. After that, these decisions work as the information source of the evidence bodies of D-S evidence theory. Finally, with the fused multimodal information and decision-making rules, diagnosis results can be given. Moreover, we describe a method to construct the basic probability assignment of evidence body from the classification result of neural network and simplify the limitation of decision rules under practical application circumstance. Experiments on several UCR standard time series datasets and the aeroengine fault datasets prove that by integrating the complementary information of multimodal neural networks which have distinctive abilities and learn sample data from different angles, a high accuracy diagnosis model can be achieved.

The rest of the paper is organized as follows. Section 2 reviews the related work on deep learning and D-S evidence

theory methods for aeroengine fault diagnosis. Section 3 presents the framework of the proposed model and the steps to implement it in details. In Section 4, experiments are carried out to validate the effectiveness of the proposed model. Finally, conclusions are drawn.

## 2. Related Work

This section will present how CNN, BPNN, and D-S evidence theory are employed for aeroengine fault diagnosis. Besides, we find that previous works ignored the diversity of aeroengine fault data, which may cause their methods hard to generalize across datasets. Referring to the shortcomings existing in previous models and the problem about datasets, we further summarize the advantages of our proposed model.

With its excellent automatic feature extraction ability, CNN success in many fields [12], including aeroengine fault diagnosis. It is used primarily to process images, videos, and other two-dimensional structured data. Then, in the field of natural language processing, convolution kernel is reshaped to one-dimensional to find the connections between words and sentences. Similarly, horizontal one-dimensional CNN is also employed for aeroengine fault diagnosis. For example, Jiang et al. [13] focused on the feature extraction ability of CNN and used it to cope with fault classification. Wang et al. [14] processed signals simply at first and then employed CNN to extract features automatically from them, avoiding the inconsistency of performance caused by traditional manual extraction of features. Some derived models of CNN have also immigrated from the initial field of image processing. As ResNet [15], it was modified into a model with supermulti-layer convolution layer to find high-level features of fault, which can obtain a promising diagnosis result. Furthermore, there are also some methods to extract the features of sample data by CNN and then use other classification methods to distinguish them. For example, in [16], the output of CNN's second-to-last layer is regarded as features of samples, and the features are transferred to SVM to be classified.

BPNN, as one of the most common and widely used artificial neural networks [17, 18], has been applied to the classification problem of aeroengine fault diagnosis long time ago [19, 20]. In past decades, researchers have tried to combine it with other methods or to optimize its structural parameters using some optimization search algorithms, in order to yield better diagnosis results. For example, in [21], to cope with the difficulty of insufficient fault samples by reducing the scale of neural network, rough set theory was used to pretreat original data, and then, BPNN was employed for fault diagnosis using the optimal decision attributes. Similarly, the method which integrated BPNN with SVM was proposed in [5]. To tackle the disadvantages of BPNN in gas turbine fault diagnosis, Yuan et al. [22] attempted to use the Particle Swarm Optimization algorithm and Levenberg Marquardt algorithm to improve the performance. Besides, BPNN is also used as a basic element of other complex models, such as methods based on nested BPNN [23] and integrated BPNN [24] for aeroengine fault diagnosis.

D-S evidence theory is one of the common means in the area of information fusion. It can integrate multiple evidence

bodies which can be the predictions of different people, data of different sensors, and results of different classifiers. Various basic models have been served as the information source of evidence bodies, for instance, RBF, SVM, and BPNN [25]; RBF and BPNN [26]; and CNN and SVM [27]. Information source of evidence bodies can be different results given by the same type of basic classifier, as well. For example, Song [28] divided the multidimensional attributes of samples into several groups and used each group to train a BPNN, creating a fusion diagnosis model of multiple BPNN. Wu et al. [29] presented an information fusion fault diagnosis method based on D-S theory using SVM. However, these means are not perfect, as they carry out D-S theory merely by incorporating homogeneous models, without taking multimodal information into account.

For aeroengine fault diagnosis, there exists a common problem: aeroengine big data has no standard format. They are diverse in quality, dimension, continuity, and order of magnitudes of the attributes. Some fault diagnosis models are sensitive to datasets and narrow in application scope. They may only have high accuracy for certain dataset. In [30], author took the influence of attribute dimensions (13 and 8, respectively) into account and found that this factor did have a great impact on the capability of the model. In [25, 31], SVM and BPNN, both as the baselines, had a conflict on whose classification accuracy is higher, because it depends on datasets. Therefore, the aeroengine fault diagnosis model cannot only be accuracy-oriented. And neural networks have strong generalization ability and certain robustness, so they can ease the problem to some extent. But some previous methods which execute D-S theory by combining deep learning models with nondeep learning models will still have that defect, since the nondeep learning model is more sensitive to datasets. This is the second point leaving to be desired for the implementation of D-S theory in field of aeroengine fault diagnosis.

Those composite models that use D-S theory have above two shortcomings, while the single modal ones which merely use CNN or BPNN do not consider information of the two excellent models at the same time, obtaining a lower diagnosis accuracy [32]. Therefore, we propose a multimodal neural network diagnosis model that learns from multiple angles. It does not only perceive features but fits the distribution of original data. Meanwhile, compared with other composite models, the proposed is composed of pure neural network. Thus, it has better robustness.

### 3. Diagnose Method Based on Multimodal Deep Neural Networks

The framework of our proposed method is shown in Figure 1. First, samples are cleaned and normalized, which is a common means to reduce the impact of various data specification. Then, CNN and BPNN are trained to the best with samples, and the scores they output are served as information sources. Finally, the D-S evidence theory is employed to coordinate the multimodal information in form of evidence bodies and generate the final decisions.

*3.1. Basic Models.* BPNN is a multilayer fully connected neural network trained by the back propagation algorithm (BP), which is established by simulating the indescribable complicated working process in the brain. Each layer in the network is composed of multiple artificial neuron cells. The weight matrix and bias in cells carry out linear transformation on its input data. Next, the transformation result is used as the net input of activation function, which then makes nonlinear mapping. Through the stacking of nonlinear mapping layer by layer and the adjustment of parameters in cells by training algorithms, BP neural network can be employed to fit complex multivariate nonlinear functions [33]. Therefore, it can be used to explore the functional or causal relationship between the overall distribution of property values and fault types.

Convolutional neural network is a feedforward neural network that consists of convolution layer, pooling layer, and fully connected layer. Compared with BPNN, CNN has characteristics of partial connection and weight sharing, which make its complexity reduces a lot. Convolution is often viewed as an effective means for feature extraction. The convolution filter used for feature extraction is a matrix in nature, which slides over input data and performs matrix operation with the data it covers, and thus, the calculation result can be regarded as the matching extent between the covered data and the feature to be extracted. Then, the result is added by bias and put into the activation function to exert a nonlinear transformation. The convolution filter is self-adaptive under the adjustment of the back propagation algorithm, so CNN has the so-called ability to automatically extract features. In the domain of aeroengine fault diagnosis, CNN is usually used to discover the features residing in attributes and establish a mapping between the fault types and the features of the properties.

The frequently used activation function mentioned before is the Rectified Linear Unit. The cross-entropy function is generally selected as the loss function for classification problems, and it can be reduced by gradient back propagation algorithm with the update of parameters in cells, as follows:

$$J(y, \hat{y} | W, b) = -y^T \log(\hat{y}), \quad (1)$$

$$\begin{aligned} \frac{\partial J(y, \hat{y})}{\partial W^{(L)}} &= \frac{\partial J(y, \hat{y})}{\partial z^{(L)}} \cdot \frac{\partial z^{(L)}}{\partial W^{(L)}} \\ \frac{\partial J(y, \hat{y})}{\partial b^{(L)}} &= \frac{\partial J(y, \hat{y})}{\partial z^{(L)}} \cdot \frac{\partial z^{(L)}}{\partial b^{(L)}}, \end{aligned} \quad (2)$$

$$\begin{aligned} W_{n+1}^{(L)} &= W_n^{(L)} - \alpha \cdot \frac{\partial J(y, \hat{y})}{\partial W_n^{(L)}} \\ b_{n+1}^{(L)} &= b_n^{(L)} - \alpha \cdot \frac{\partial J(y, \hat{y})}{\partial b_n^{(L)}}, \end{aligned} \quad (3)$$

Herein,  $J(\cdot)$  represents the loss function.  $W$  and  $b$  are parameters in cells, representing the weight matrix and bias, respectively.  $y$  and  $\hat{y}$  represent the real and predicted categories of sample, expressed in one-hot vector form.

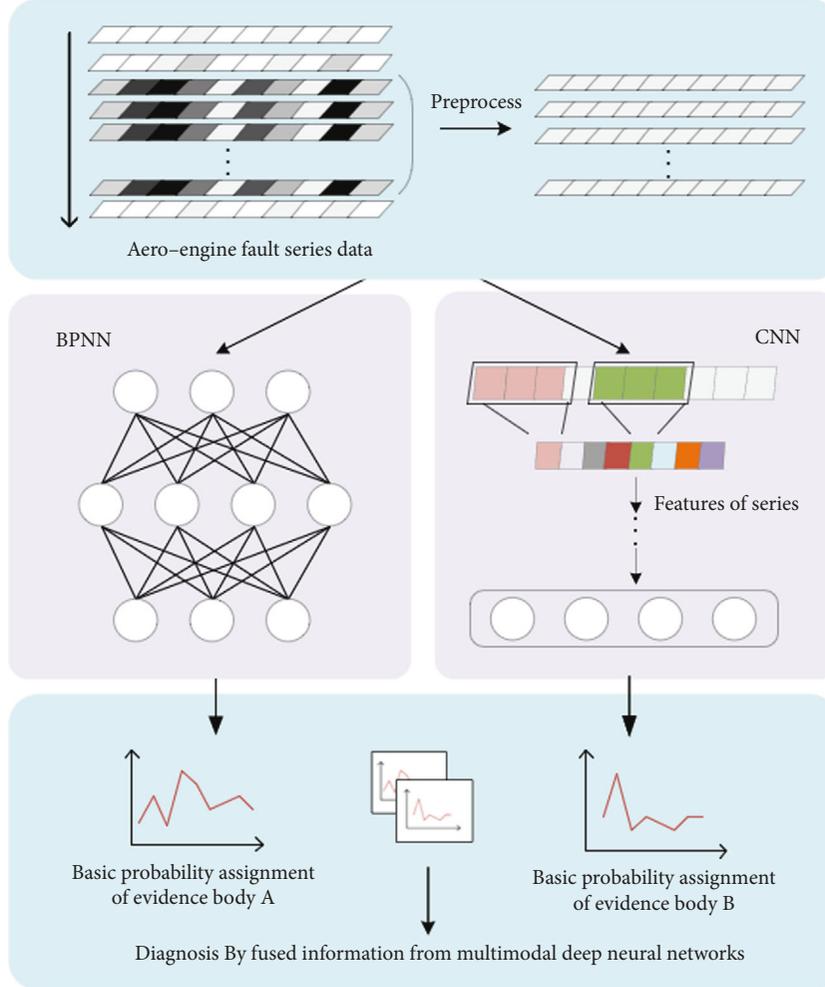


FIGURE 1: The architecture of multimodal deep neural networks diagnosis model.

The subscript  $n$  represents the parameter after the  $n$ -th training, and  $L$  represents the  $L$ -th layer.  $Z = W^T x + b$  is the net input of activation function.  $\alpha$  is the learning rate.

D-S evidence theory, another basic algorithm used in this paper, roughly includes three phases: constructing the basic probability assignment of each evidence body, fusion of evidence bodies, and giving final result according to the decision rules. In D-S evidence theory, the conclusion given by each evidence body about the hypothetic categories that samples may belong to is presented as a probability assignment. Since the decision information of each evidence body is not authoritative, the uncertainty of decision should be considered when constructing basic probability assignment from information source. In fusion phase, the basic probability assignments of different evidence bodies will be combined into a new fused one as output. Then, according to the decision rules, final classification results can be obtained. In theory, only when the difference between the maximum and the second largest values of the fused probability assignment is greater than a certain threshold, can the final classification decision be considered as an effective judgment, since the uncertainty information is involved in the probability

calculation of every hypothesis. Otherwise, if they are close to each other, the given decision may be a misjudgment caused by uncertain information.

Many researchers have been engaged in the improvement of D-S theory, including how to scientifically construct the evidence bodies, improve the fusion formulas, and solve the problem of evidence conflict. Consequently, various specific formulas and rules to implement D-S theory are constructed. When applied, the specific implementary formulas used by different authors and in different scenes may vary greatly. In this paper, we design a new D-S theory-based multimodal fusion method.

**3.2. The Proposed Method.** Steps and details of the proposed model are as follows:

The first step is to preprocess the aeroengine sequence data, so that they can be directly input into the neural networks and can be better applied for feature perception and data fitting. Specifically, ignore the sequences with too many zero values at the beginning and end of datasets, and normalize the remaining samples by attributes according to formula (4):

$$X_{ij} = \frac{X_{ij} - \min \{X_{1j}, X_{2j}, \dots, X_{Nj}\}}{\max \{X_{1j}, X_{2j}, \dots, X_{Nj}\} - \min \{X_{1j}, X_{2j}, \dots, X_{Nj}\}} \quad (4)$$

$(i = 1, 2, \dots, N; j = 1, 2, \dots, d)$

where  $X_{ij}$  represents the  $j$ -th attribute of the  $i$ -th sequence data in samples.  $N$  is the number of sequence data in samples.  $d$  represents the dimension of attributes.  $\max \{\cdot\}$  means the maximum value and  $\min \{\cdot\}$  means the minimum value.

Labels should be added one by one to samples with  $1, 2, \dots, m$  representing fault types  $F_1, F_2, \dots, F_m$ . After that, merge and randomly shuffle all fault sequence data. On the one hand, for the problem of sequence classification, CNN and BPNN do not depend on the order relationship between samples. On the other hand, the shuffle can increase the randomness of samples, then train favorable models. At last, samples are divided into training set  $D$ , validation set  $V$ , and test set  $T$ .

The second step is to train CNN and BPNN. Thus, these two trained models are employed to diagnose the samples in validation set  $V$  and test dataset  $T$ , respectively. The scores they given before *SoftMax* will be used as the information source in the next step.

Generally, the number of fully connected layers in neural network should be less than or equal to three. Regarding the fault diagnosis of aeroengine, two or three layers should be set due to the complicated relationship between input data and types of faults. For time series data like aeroengine data, one-dimensional convolution should be carried out, and the size of convolution kernel should be smaller than the dimension of attributes. The training process uses the cross-entropy loss function and the back propagation algorithm given in formulas (1)–(3).

The third step is to convert multimodal information into the basic probability assignment of evidence bodies, with the selected coordination factor  $\rho$ , which will be discussed in the next subsection 3.3.

For each sequence data sample, its scores of belonging to different fault types before *SoftMax* given by BPNN and CNN should be converted into the initial probability assignments of evidence bodies. These can be calculated as

$$E_i^{M0}(l) = \text{softmax}(\rho \cdot r_i(l)) = \frac{\exp(\rho \cdot r_i(l))}{\sum_{l=1}^m \exp(\rho \cdot r_i(l))}. \quad (5)$$

Herein,  $r$  represents the scores before *SoftMax* function, given by the neural network.  $\rho$  is the coordination factor of information source.  $E_i^{M0}$  represents the initial probability assignment of samples. The subscript  $i$  represents the  $i$ -th sequence data of samples. Superscript  $M$  represents an evidence body, denoted by  $M = \{A : CNN, B : BPNN\}$ . And  $l = 1, 2, 3, \dots, m$  refers to one hypothesis of the fault categories.

Thus, the initial probability assignment can be used to calculate the uncertainty measurement of samples, according to formulas (6) and (7). The sum of squares of the distance

from the initial probability assignment of one evidence body to the average initial probability assignment of two evidence bodies is selected as the uncertainty measurement. This value measures the conflict between decisions for the same sample data given by two evidence bodies. The larger it is, the more inconsistent the two judgments are, which means the less determinate they are.

$$AVG\_E_i(l) = \frac{1}{2} [E_i^{A0}(l) + E_i^{B0}(l)], \quad (6)$$

$$D_i^M = \sum_{l=1}^m [E_i^{M0}(l) - AVG\_E_i(l)]^2. \quad (7)$$

In the formulas,  $AVG\_E$  represents the average of the two evidence bodies.  $E^{A0}$  and  $E^{B0}$  represents the initial probability assignments of the two evidence bodies.  $D_i^M$  represents the sum of the squares of the distance from the result of the evidence body  $M$  to the average result. Then, the measurement can be converted into the uncertainty of the sample using formula (8). Thus, the basic probability assignment of evidence body can be calculated by formula (9) and (10), with initial probability assignment and uncertainty:

$$U_i^M = \sqrt{D_i^M}, \quad (8)$$

$$E_i^M(l) = E_i^{M0}(l) \times (1 - U_i^M), \quad (9)$$

$$E_i^M(\Theta) = U_i^M, \quad (10)$$

where  $E_i^M(l)$  represents the belief probability of hypothesis  $l$  in basic probability assignment.  $E_i^M(\Theta)$  represents the probability of belonging to the item of universal set  $\Theta$ .  $U_i^M$  represents the uncertainty of sample under evidence body  $M$ .

The fourth step is to determine how to fuse the multimodal information which are expressed in form of evidence bodies and give the decision-making criteria.

The basic probability assignments of CNN and BPNN are used to calculate their counterpart after fusion, as shown in formulas ((11) and ((12):

$$E_i(l) = \frac{\sum_{k,h} E_i^A(k) \times E_i^B(h)}{1 - U_i} A_i(k) \cap B_i(h) = l, \quad (11)$$

$$U_i = \sum_{k,h} E_i^A(k) \times E_i^B(h) A_i(k) \cap B_i(h) = \emptyset \text{ or } \Theta. \quad (12)$$

Herein,  $k, h \in \{F_1, F_2 \dots F_m\} \cup \{\Theta\}$  represent two hypotheses of assignment.  $A_i(k), B_i(h)$  represent the hypotheses taken, respectively, from two evidence bodies for the same sequence.  $U_i$  represents the sample's uncertainty of assignment after fusion.  $E_i(l)$  represents the belief that this sequence belongs to hypotheses  $l$ , according to the fused probability assignment.

Diagnose Method Based on Multimodal Deep Neural Networks.

**Input:** Aero engine gas-path fault datasets of  $m$  fault modes.

**Output:** Diagnosis result  $Y$  of sample sequence in test set  $T$

```

1  Preprocess the input dataset: clean, normalize by formula (4), label, merge and shuffle it. Divide it into training set  $D$ , validation set  $V$  and test set  $T$ .
2  For  $M$  in {A: CNN, B: BPNN}:
3    Repeat:
4      Adjust one hyper-parameter, including layers, learning rate, etc.
5      For  $i < \text{maxiterations}$ :
6        Calculate the value of cross-entropy loss.
7        Update parameters by BP algorithm:  $W_i^{(l)} \leftarrow W_{i-1}^{(l)}, b_i^{(l)} \leftarrow b_{i-1}^{(l)}$ .
8      End For
9    Until: Performance of  $M$  will no longer be improved.
10   Give the diagnosis scores  $r$  of  $V$  and  $T$ , by  $M$  before its SoftMax process.
11 End For
12 Choose one scores information to be adjust, and fix the other  $\rho=1$ .
13 Repeat
14   Adjust  $\rho$  of the selected scores information.
15   Execute steps given in 20-23 on validation set  $V$ .
16   Calculate  $std(E_i)$  of its initial probability assignment by formula (14).
17 Until: Find the continuous range of  $\rho$  where diagnosis result for  $V$  are best.
18 Select a value for  $\rho$  with proper standard deviation in that range.
19 For  $M$  in {A: CNN, B: BPNN}:
20   With the selected  $\rho$ , compute basic probability assignments by formula (5) to (10):
       $E_i^{M0}(l) = softmax(\rho \cdot r_i(l))$ 
       $D_i^M \leftarrow E_i^{A0}(l), E_i^{B0}(l)$ 
       $E_i^M(l), E_i^M(\Theta) \leftarrow E_i^{M0}(l), D_i^M$ 
21 End For
22 Fusion the multimodal evidence bodies by formula ((11) and ((12):
 $E_i(l) \leftarrow E_i^A(k), E_i^B(h)$ 
23 According to the comprehensive information after fusion, give decision by rule:
 $Y_i = argMax(E_i(l))$ 
24 Return:  $Y_i$ 

```

ALGORITHM.

Therefore, the decision rule can be achieved by formula (13). For each sequence, the hypothesis with the highest probability is selected as the final result.

$$Y_i = \arg\text{Max}(E_i(l)), \quad (13)$$

where  $Y_i$  is the diagnosis result of the  $i$ -th sequence in datasets.

In order to give each sample a definite category, the final probability assignment can no longer exist the universal set item  $\Theta$ , which represents uncertain category. Besides, according to D-S theory, the difference of belief among hypotheses should be considered when formulating decision rules. Whereas, for the same purpose, no threshold limitation is applied in our decision-making criteria.

The details of the proposed method are shown in the Algorithm.

**3.3. Factor Selection.** The complementary multimodal information for integration should be at the same crucial level. Indeed, this cannot be guaranteed. If one neural network gives scores of samples all like “0.99-0.005-0.005-0,” while the other one is “0.85-0.10-0.03-0.02” (take the assignment

TABLE 1: Details of the UCR standard datasets.

| Dataset       | Sequence amount | Attributes number | Classes number |
|---------------|-----------------|-------------------|----------------|
| Chlor.Conc    | 4307            | 166               | 3              |
| Cinc_ECG      | 1420            | 1639              | 4              |
| Dist.phal.O.C | 876             | 80                | 2              |
| ECGFivedays   | 884             | 136               | 2              |
| Yoga          | 3300            | 426               | 2              |

with four hypotheses as an example), it will inevitably prevent the second information source from playing its role. So, the adjustment is necessary, and the coordination factor  $\rho$  is thus brought in.

The adjustment needs to be done on just one of the two output information from CNN and BPNN, while the other is fixed as  $\rho = 1$ . The value of factor  $\rho$  can be determined according to the accuracy variation of the validation set  $V$ . And the relative magnitudes of standard deviations of two initial probability assignments  $std(E_i)$  can serve as a reference index, as well.

TABLE 2: Comparison of experimental results on UCR standard datasets.

| Dataset       | Proposed | NN-based methods |       |       |       |       | Distance-based methods |       |       |
|---------------|----------|------------------|-------|-------|-------|-------|------------------------|-------|-------|
|               |          | ResNet           | MCNN  | CTN-T | ESN   | TN    | ST                     | LWDTW | BOSS  |
| Chlor.Conc    | 0.999    | 0.84             | 0.797 | 0.83  | 0.920 | 0.731 | 0.700                  | 0.644 | 0.66  |
| Cinc_ECG      | 1.00     | —                | 0.942 | —     | 0.679 | —     | 0.846                  | 0.935 | 0.901 |
| Dist.phal.O.C | 0.864    | 0.80             | —     | 0.80  | —     | 0.812 | —                      | 0.739 | 0.815 |
| ECGFivedays   | 1.00     | 0.97             | 1.00  | 1.00  | 1.00  | 0.926 | 0.999                  | 0.835 | 0.983 |
| yoga          | 0.941    | 0.87             | 0.888 | 0.92  | 0.820 | 0.84  | 0.846                  | 0.847 | 0.901 |
| average       | 0.961    | 0.88             | 0.907 | 0.91  | 0.855 | 0.827 | 0.838                  | 0.800 | 0.852 |

For each value of the undetermined  $\rho$ , a diagnosis accuracy of  $V$  can be obtained by processing its scores information using formula (5)–(13). And  $std(E_i)$  should be calculated at the same time with the following formula.

$$std(E_i) = \left( \frac{1}{m} \sum_l \left( E_i(l) - \frac{1}{m} \right)^2 \right)^{1/2}, \quad (14)$$

When the coordination factor  $\rho$  within a certain continuous range, the diagnostic accuracy of  $V$  can reach the highest. The proper value of the undetermined  $\rho$  should generate in this range. Meanwhile,  $std(E_i)$  of the two evidence bodies under the selected value should be similar. More directly, their ratio should be approximately between 0.9 and 1.1.

#### 4. Experiments

Since there is no available public standard dataset for aeroengine fault diagnosis, and most artificial intelligence methods of this problem have not opened their source codes, it is difficult to compare the performance of the proposed model with other methods. Fortunately, aeroengine fault data can be treated as time series data, and most methods used to classify time series data can be employed for aeroengine fault diagnosis. The effectiveness of the proposed model can be validated by comparison experiments on standard datasets of UCR (University of California, Riverside) time series.

Therefore, the experimental verification can be done in two parts. Firstly, the proposed method is used to classify standard UCR time series datasets and compared with other time series classification models. Secondly, its effectiveness is verified in the target field by using the real aeroengine fault datasets.

*4.1. Experiments on Standard UCR Time Series Datasets.* Experiments are firstly executed on several UCR standard datasets, including Chlor.Conc, Cinc\_ECG, Dist.phal.O.C, ECGFivedays, and yoga. Details of them are shown in Table 1.

The comparisons among diagnosis accuracies of various models are shown in Table 2, and the specific experiment results of the proposed could be found in Table 3. CTN-T is a deep learning method based on single modal CNN and applying the transfer learning idea [34]. MCNN take into account the features at different time scales. It can be

TABLE 3: Comparison of the performance without and with  $\rho$ .

| Dataset       | BPNN   | CNN    | Multimodal results |        |
|---------------|--------|--------|--------------------|--------|
|               |        |        | Without            | With   |
| Chlor.Conc    | 0.991  | 0.993  | 0.999              | 0.999  |
| Cinc_ECG      | 0.958  | 0.996  | 1.00               | 1.00   |
| Dist.phal.O.C | 0.835  | 0.858  | 0.858              | 0.864  |
| ECGFivedays   | 0.983  | 1.00   | 1.00               | 1.00   |
| yoga          | 0.931  | 0.926  | 0.939              | 0.941  |
| Type A        | 0.9598 | 0.9489 | 0.9621             | 0.9627 |
| Type B        | 0.9818 | 0.9771 | 0.9843             | 0.9864 |

regarded as a model which ensemble multiple CNN at feature level [35]. ResNet is a typical neural network, and paper [36] conducted an experiment to test its ability to classify time series data. LW-DTW is k-nearest-neighbors approach based on a locally weighted dynamic time warping [37]. BOSS proposes a distance which is based on histograms of symbolic Fourier approximation words [38]. The results are from University of East Anglia website. Libsvm library (2017). <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>. ST (Shapelet Transform) classifies samples by considering the similarity between a shapelet and a sequence, where shapelet is a time-series subsequence. Herein, LW-DTW and CTN-T are the latest methods for time series classification proposed by researchers in 2019 and have good performance. Since the data includes temporal information, comparisons to RNN-based method are also taken into consideration, though classification methods on RNN in recent years are quite few. TN (TimeNet) [39] is a generic off-the-shelf feature extractor for time series data based on RNN automatic encoder, which can be combined with SVM and other classifiers. ESN (Echo state network) is a form of RNN. The authors tried classical and improved ESN-based approach and their fusion version [40]. We took the best results of them for comparison. Besides, RNN-based models are rarely used alone, so in subsequent analyses, they do not perform as representations of deep learning models in this field.

Horizontal contrast shows that the performance of deep learning models is better than distance-based methods on the whole. The overall accuracy of ResNet, MCNN, and CTN-T on the above datasets is about 90%, higher than that of Shapelet, LW-DTW, and BOSS methods. Distance-based methods distinguished samples by its closeness, which does

| Basic probability assignment of CNN |       |       |       |       | Basic probability assignment of BPNN |       |       |       |       | Fused information |       |       |       | Label |       |       |   |
|-------------------------------------|-------|-------|-------|-------|--------------------------------------|-------|-------|-------|-------|-------------------|-------|-------|-------|-------|-------|-------|---|
| 0.001                               | 0.084 | 0.102 | 0.184 | 0.188 | 0.441                                | 0.000 | 0.002 | 0.315 | 0.000 | 0.019             | 0.664 | 0.001 | 0.057 | 0.271 | 0.122 | 0.140 | 3 |
| 0.001                               | 0.003 | 0.479 | 0.007 | 0.418 | 0.092                                | 0.000 | 0.000 | 0.281 | 0.000 | 0.415             | 0.304 | 0.000 | 0.001 | 0.441 | 0.002 | 0.512 | 5 |
| 0.129                               | 0.004 | 0.006 | 0.016 | 0.387 | 0.457                                | 0.246 | 0.077 | 0.000 | 0.001 | 0.000             | 0.676 | 0.264 | 0.038 | 0.004 | 0.011 | 0.262 | 1 |
| 0.000                               | 0.158 | 0.003 | 0.000 | 0.366 | 0.472                                | 0.000 | 0.300 | 0.006 | 0.001 | 0.006             | 0.687 | 0.000 | 0.345 | 0.005 | 0.001 | 0.259 | 2 |
| 0.240                               | 0.232 | 0.022 | 0.227 | 0.001 | 0.279                                | 0.085 | 0.021 | 0.000 | 0.365 | 0.000             | 0.528 | 0.191 | 0.138 | 0.011 | 0.387 | 0.000 | 4 |
| 0.001                               | 0.011 | 0.548 | 0.001 | 0.247 | 0.192                                | 0.000 | 0.000 | 0.232 | 0.000 | 0.329             | 0.439 | 0.001 | 0.005 | 0.540 | 0.000 | 0.334 | 3 |
| 0.315                               | 0.007 | 0.000 | 0.000 | 0.000 | 0.677                                | 0.003 | 0.174 | 0.000 | 0.000 | 0.000             | 0.823 | 0.264 | 0.126 | 0.000 | 0.000 | 0.000 | 1 |
| 0.108                               | 0.285 | 0.066 | 0.261 | 0.009 | 0.272                                | 0.262 | 0.215 | 0.000 | 0.002 | 0.000             | 0.521 | 0.183 | 0.329 | 0.034 | 0.138 | 0.004 | 2 |
| 0.000                               | 0.000 | 0.057 | 0.000 | 0.624 | 0.319                                | 0.000 | 0.000 | 0.233 | 0.000 | 0.203             | 0.564 | 0.000 | 0.000 | 0.132 | 0.000 | 0.671 | 5 |
| 0.010                               | 0.207 | 0.001 | 0.461 | 0.003 | 0.318                                | 0.000 | 0.333 | 0.000 | 0.103 | 0.000             | 0.564 | 0.005 | 0.360 | 0.001 | 0.388 | 0.002 | 4 |

FIGURE 2: Basic probability assignment heat map of examples corrected by multimodal information.

not reflect the distribution of sample's attributes. While deep learning methods fitted the samples intensively; thus, they can better cope with the highly complex time series data. On the above dataset (some details could be found in Table 3), the basic deep neural network model like CNN and BPNN can achieve higher diagnosis accuracy than other complex versions. And after fusion, better performance can be obtained. Even in the case that the accuracy of single modal model like CNN reaches extreme high of 99%, for instance, on datasets Chlor.Conc and Cinc\_ECG, it can still be improved in fused model. It was the comprehensive diagnosis information that contribute to the superior result. Which lead the multimodal model to have capabilities of BPNN and CNN at the same time. Figure 2 in the next experiments will give a more intuitive and detailed explanation for this. In addition, another support for this interpretation is that MCNN, the integration of homogeneous model, is even not as accurate as the single modal fully connected neural network on many datasets, in the comparison of the original paper [35].

**4.2. Experiment on Real Aeroengine Fault Dataset.** The aeroengine fault datasets used in this experiment are from AVIC Shenyang Engine Design Institute (606 institute). These datasets contain fault data of engine type A and type B, and the data has 23 dimensional attributes and 5 fault modes. Through experimental comparison, each model can be best trained when total 60,000 samples of various fault modes are taken. Samples were divided into training set, validation set, and test set according to the proportion of 0.7, 0.1, and 0.2. In addition, the collection surrounding of above datasets is a laboratory with relatively stable temperature and pressure. In order to simulate the influence brought by terrible environment at high altitude that aeroengines work, a quarter of sample data are randomly selected and added with Gaussian white noise.

Heat map of basic probability assignment in Figure 2 shows how sequences are corrected by multimodal information. Where the sixth term in basic probability assignment of BPNN and CNN represents its uncertainty. The first five sequences are corrected by the contribution of BPNN, and the last five are by that of CNN. Due to CNN and BPNN focus on different aspects of samples, they might give inconsistent diagnosis result for the same sequence. For the case where one model gives right result while the other judge

TABLE 4: Experiment results on aeroengine fault datasets.

| Dataset           | CNN    | BPNN   | Multimodal NN |
|-------------------|--------|--------|---------------|
| Aeroengine type A | 0.9598 | 0.9489 | 0.9627        |
| Aeroengine type B | 0.9818 | 0.9771 | 0.9864        |

wrongly, the fusion of multimodal information may lead to the correct final result. Therefore, on the whole, the multimodal model can correctly diagnose more samples.

Table 4 shows the diagnosis accuracy of two kinds of single modal neural networks and that of multimodal one. On the two types aeroengine fault datasets, the accuracies of the multimodal neural network are improved by 0.29% and 0.46%, compared with the maximum values between the basic CNN and BPNN. Observation shows that the fusion of multimodal decision information by D-S evidence theory can effectively improve the diagnosis accuracy, and this model has a good performance on the problem of aeroengine fault diagnosis.

Meanwhile, we also paid attention to the performance improvement brought by the coordination factor  $\rho$  during the experiment. Except for the three datasets with extremely high diagnosis accuracy given by the basic model, it plays a good role on both other UCR datasets and aeroengine fault datasets. Particularly, for the dataset Dist.Phal.O.C, the coordination factor  $\rho$  makes the contribution of evidence theory improved from none to increasing accuracy by 0.6%. It enables the evidence theory to better integrate the multimodal information.

## 5. Conclusion

This paper proposes a multimodal aeroengine fault diagnosis model based on CNN and BPNN, aiming at obtaining high diagnosis accuracy and relatively strong robustness by comprehensively considering the complementary decision information from pure neural network. Firstly, sequence data will be preprocessed. Then, the two neural networks are trained with it. And then, the multimodal information of two networks are fused by D-S evidence theory. At last, final diagnosis results are given. This paper also gives a method to construct the basic probability assignment of evidence body by the output scores information of neural network, as well

as the fusion formulas and decision rules of D-S theory for practical application.

The two modal neural networks, BPNN and CNN, have the same level but distinctive ability. BPNN can establish mapping between a certain type of fault and the overall distribution of all attributes in sample, and CNN using one-dimensional convolution kernels can extract features from samples and explore relationship between features of attributes and the existence of faults. Moreover, pure neural network-based method can better adapt to input data of various specifications.

Experiment on UCR standard time series datasets proves that classification methods based on neural network are better than methods based on distance, and the proposed model has better performance than some typical models and latest methods on certain datasets. And though ultrahigh accuracy the single modal neural network has, it can still be enhanced by fusing multimodal diagnosis results. Likewise, the experiment conducted on the aeroengine fault data has also achieved good results, which illustrates that neural network does have a strong ability to cope with aeroengine fault diagnosis, and the fusion of multimodal information at decision level is indeed an effective way to further improve accuracy.

## Data Availability

The .mat data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (61906030) and the Science and the Fundamental Research Funds for the Central Universities (DUT20RC(4)009).

## References

- [1] Q. Pan, Y. Liu, R. Zhou, H. Wang, H. Chen, and T. He, "An automatic abrupt signal extraction method for fault diagnosis of aero-engines," *Journal of Mechanical Science and Technology*, vol. 33, no. 4, pp. 1633–1640, 2019.
- [2] Q. Huang, G. Zhang, T. Zhang, and J. Wang, "A kind of approach for aero engine gas path fault diagnosis," in *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, Location: Dallas, TX, USA, 2017.
- [3] Z. B. Shi, Q. G. Song, M. Z. Ma, and Q. Li, "Fault diagnosis of steam turbine based on mpso-svm algorithm," *Journal of Chinese Society of Power Engineering*, vol. 32, no. 6, pp. 454–462, 2012.
- [4] X. Shou, "Aero-engine fault diagnosis based on support vector machine," *Mechanical Science and Technology*, 2005.
- [5] D. Seo, T. Roh, and D. Choi, "Defect diagnostics of gas turbine engine using hybrid SVM-ANN with module system in off-design condition," *Journal of Mechanical Science and Technology*, vol. 23, no. 3, pp. 677–685, 2009.
- [6] F. Xia, H. Zhang, D. Peng, H. Li, and Y. Su, "Turbine Fault Diagnosis Based on Fuzzy Theory and SVM," in *Artificial Intelligence and Computational Intelligence*, vol. 5855, Springer, Berlin, Heidelberg, 2009.
- [7] C. Ren, H. Dong, P. Hou, X. Dong, and Y. Tao, "A clustering-based method for health conditions evaluation of aero-engines," in *2019 Prognostics and System Health Management Conference (PHM-Paris)*, Paris, France, France, 2019.
- [8] X. S. Lin, B. W. Li, and X. Y. Yang, "Engine components fault diagnosis using an improved method of deep belief networks," in *2016 7th International Conference on Mechanical and Aerospace Engineering (ICMAE)*, London, UK, 2016.
- [9] R. Jiang and W. Zhu, "A pnn fault diagnosis method for gas turbine," in *World Automation Congress 2012*, pp. 1–4, Puerto Vallarta, Mexico, Mexico, 2012.
- [10] Z. Ning, R. Y. K. Kwok, K. Zhang et al., "Joint computing and caching in 5g-envisioned internet of vehicles: A deep reinforcement Learning-Based traffic control system," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–12, 2020.
- [11] L. Zhao, Z. Chen, L. T. Yang, M. J. Deen, and Z. J. Wang, "Deep semantic mapping for heterogeneous multimedia transfer learning using co-occurrence data," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 15, no. 1s, pp. 1–21, 2019.
- [12] Z. Ning, Y. Li, P. Dong et al., "When deep reinforcement learning meets 5g-enabled vehicular networks: a distributed offloading framework for traffic big data," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 1352–1361, 2020.
- [13] Z. Jiang, H. Fang, H. Shi, S. Ren, H. Yang, and F. Wang, "The prognostic method of engine gas path based-on convolutional neural network," *DEStech Transactions on Computer Science and Engineering*, no. iciti, 2019.
- [14] J. Wang, J. Zhuang, L. Duan, and W. Cheng, "A multi-scale convolution neural network for featureless fault diagnosis," in *2016 International Symposium on Flexible Automation (ISFA)*, pp. 65–70, Cleveland, OH, USA, 2016.
- [15] L. Wen, X. Li, and L. Gao, "A transfer convolutional neural network for fault diagnosis based on resnet-50," *Neural Computing and Applications*, vol. 32, no. 10, pp. 6111–6124, 2020.
- [16] S. Zhong, S. Fu, and L. Lin, "A novel gas turbine fault diagnosis method based on transfer learning with CNN," *Measurement*, vol. 137, pp. 435–453, 2019.
- [17] X. Wang, Z. Ning, S. Guo, and L. Wang, "Imitation learning enabled task scheduling for online vehicular edge computing," *IEEE Transactions on Mobile Computing*, 2020.
- [18] Z. Ning, K. Zhang, X. Wang et al., "Intelligent edge computing in internet of vehicles: a joint computation offloading and caching solution," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–14, 2020.
- [19] C. Angelakis, E. N. Loukis, A. Pouliezios, and G. S. Stavrakakis, "A neural network-based method for gas turbine blading fault diagnosis," *International Journal of Modelling and Simulation*, vol. 21, no. 1, pp. 51–60, 2001.
- [20] S. O. T. Ogaji and R. Singh, "Advanced engine diagnostics using artificial neural networks," *Applied Soft Computing*, vol. 3, no. 3, pp. 259–271, 2003.
- [21] N. Zhao, S. Y. Li, S. Yi, Y. P. Cao, and Z. T. Wang, "Fault diagnosis based on rough set and bp neural network (rs-bp) for gas turbine engine," *Advanced Materials Research*, vol. 732-733, pp. 397–401, 2013.

- [22] B. Yuan, F. Xia, Z. Wang, and H. Tie, "A comparative research based on three different algorithms for fault diagnosis in gas turbine," in *2017 4th International Conference on Systems and Informatics (ICSAI)*, Hangzhou, China, 2017.
- [23] A. D. Fentaye, A. T. Baheta, and S. I. Gilani, "Gas turbine gas-path fault identification using nested artificial neural networks," *Aircraft Engineering and Aerospace Technology*, vol. 90, no. 6, pp. 992–999, 2018.
- [24] S. Xiangyang, "Research on aero-engine fault diagnosis based on integrated neural network," *Mathematical Models in Engineering*, vol. 5, no. 2, pp. 41–47, 2019.
- [25] Y. Xiaohong, G. Haifeng, Z. Jing, X. Jing, and Z. Dandan, "Aero engine gas path fault prediction based on multi-sensor information fusion," in *2016 IEEE Chinese Guidance, Navigation and Control Conference (CGNCC)*, Nanjing, China, 2016.
- [26] C. Xu, H. Zhang, D. Peng, Y. Yu, C. Xu, and H. Zhang, "Study of fault diagnosis of integrate of D-S evidence theory based on neural network for turbine," *Energy Procedia*, vol. 16, no. 2012, pp. 2027–2032, 2012.
- [27] L. Xiaolin and Q. Weidong, "Aero-engine fault fusion diagnosis based on D-S evidence theory," in *Proceedings of the 37th Chinese Control Conference*, pp. 963–967, Wuhan, China, 2018.
- [28] H. Song, *Aeroengine fault diagnosis based on information fusion technology*, [Ph. D. Thesis], Central South University, 2013.
- [29] W. J. Wu, D. G. Huang, and Z. Dong, "Research on a fault diagnosis method for aero-engine based on improved svm and information fusion," *Applied Mechanics and Materials*, vol. 66-68, pp. 811–816, 2011.
- [30] G. Lan, N. Cheng, and Q. Li, "Comparison and fusion of various classification methods applied to aero-engine fault diagnosis," in *2017 29th Chinese Control And Decision Conference (CCDC)*, Chongqing, China, 2017.
- [31] P. Jun and H. Jiangbo, "Aero-engine fault diagnosis based on ipso-elman neural network," *Journal of Aerospace Power*, vol. 32, no. 12, pp. 3031–3038, 2017.
- [32] L. Zhao, T. Yang, J. Zhang, Z. Chen, Y. Yang, and Z. J. Wang, "Co-learning non-negative correlated and uncorrelated features for multi-view data," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–11, 2020.
- [33] Z. Ning, P. Dong, X. Wang et al., "Mobile Edge Computing Enabled 5g Health Monitoring for Internet of Medical Things: A Decentralized Game Theoretic Approach," *IEEE Journal on Selected Areas in Communications*, to appear, 2020.
- [34] K. Kashiparekh, J. Narwariya, P. Malhotra, L. Vig, and G. Shroff, "ConvtimeNet: A pre-trained deep convolutional neural network for time series classification," in *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, Hungary, 2019.
- [35] Z. Cui, W. Chen, and Y. Chen, "Multi-scale convolutional neural networks for time series classification Computer Vision and Pattern Recognition," 2016, <https://arxiv.org/abs/1603.06995>.
- [36] Z. Wang, W. Yan, and T. Oates, "Time series classification from scratch with deep neural networks: A strong baseline," in *2017 International Joint Conference on Neural Networks (IJCNN)*, Anchorage, AK, USA, 2017.
- [37] J. Yuan, A. Douzal-Chouakria, S. Varasteh Yazdi, and Z. Wang, "A large margin time series nearest neighbour classification under locally weighted time warps," *Knowledge and Information Systems*, vol. 59, no. 1, pp. 117–135, 2019.
- [38] P. Schafer, "The boss is concerned with time series classification in the presence of noise," *Data Mining and Knowledge Discovery*, vol. 29, no. 6, pp. 1505–1530, 2015.
- [39] P. Malhotra, V. Tv, L. Vig, P. Agarwal, and G. Shroff, "Timenet: Pre-Trained Deep Recurrent Neural Network for Time Series Classification: Learning," 2017, <https://arxiv.org/abs/1706.08838>.
- [40] W. Aswolinskiy, R. F. Reinhart, and J. Steil, "Time series classification in reservoir- and model-space," *Neural Processing Letters*, vol. 48, no. 2, pp. 789–809, 2018.