

Fault detection and diagnosis for Air Handling Unit based on multiscale convolutional neural networks

Fanyong Cheng ^{a,b,d}, Wenjian Cai ^b, Xin Zhang ^{c,e,*}, Huanyue Liao ^b, Can Cui ^b

^a Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, School of Electrical Engineering, Anhui Polytechnic University, Wuhu, 241000, China

^b School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

^c College of Electrical Engineering, Zhejiang University, Hangzhou 310058, China

^d Anhui Key Laboratory of Detection Technology and Energy Saving Devices, Anhui Polytechnic University, Wuhu, 241000, China

^e Hangzhou Global Scientific and Technological Innovation Center, Zhejiang University, Hangzhou 310058, China

ARTICLE INFO

Article history:

Received 11 August 2020

Revised 18 January 2021

Accepted 26 January 2021

Available online 1 February 2021

Keywords:

Convolutional neural networks

Multiscale signals

HVAC system

Air Handling Unit

Fault diagnosis

ABSTRACT

This paper proposes a novel fault detection and diagnosis (FDD) method using multiscale convolutional neural networks (MCNNs) for Air Handling Unit (AHU) in Heating, Ventilation, and Air conditioning (HVAC) system. In existing works, it is challenging to achieve high diagnosis performance on multiscale monitoring signals from AHU system since the feature extraction methods in these works are not powerful enough. Although the single-scale convolutional neural networks (CNNs) have been adopted to improve the ability of feature extraction in FDD, it remains difficult to extract strong discriminative feature from multiscale monitoring signals only using single-scale kernels. In this paper, a novel MCNNs-based FDD method is proposed with three different scale kernels to improve the ability of feature extraction and the end-to-end learning strategy is adopted to optimize the model of MCNNs. With strong representation ability, the proposed method can capture highly discriminative features, which can help to improve the diagnostic performance of AHU. The proposed method is compared with other five commonly used methods using the measured data from our AHU experiment platform. The comparison results demonstrate that the proposed MCNNs-based FDD method outperforms other FDD methods.

© 2021 Elsevier B.V. All rights reserved.

1. Introduction

Heating, Ventilation, and Air conditioning (HVAC) system is used to provide healthy and comfortable indoor conditions for occupants [1,2]. In the HVAC system, Air Handling Unit (AHU) is the most extensively operated component to improve comfort of living environment and efficiency of energy usage [3]. It includes a considerable number of sensors, control actuators, coils, filters, and dampers, etc. These components are at risk of failure, which may result in reduced efficiency and equipment deterioration, and shorten using life of other related components. Also, faults would cause air quality degradation affecting the health of living. Hence, fault detection and diagnosis (FDD) of AHU is essential to keep comfort and energy-saving and provide preventive maintenance for the HVAC system.

Various solutions of FDD [4,5] in the HVAC system have been investigated and proposed to prevent faults resulting in thermal

discomfort and energy inefficiency in buildings. Among them, many studies have focused on developing FDD methods for AHU. These methods can be categorized as model-based methods [6,7], data-driven methods [8–11]. Model-based FDD methods are developed to design the rules for FDD depending on adequate prior knowledge concerning the HVAC system. This kind of diagnosis method is not general for different systems and hard to accurately detect complicated faults under noise. Due to the limited performance of model-based method, many data-driven methods had been adopted to design FDD methods for AHU systems, such as principal component analysis (PCA) [8,9], wavelet transform (WT) [10], extreme learning machine (ELM) [11], Support Vector Machine (SVM) [12] and neural networks (NNs) [13–15]. These methods are difficult to obtain high diagnosis performance due to the limitation of feature representation. Therefore, a fault diagnostic model based on deep neural networks was proposed for AHU [16] and obtained 95.16% accuracy. This deep model is difficult to capture multiscale information for higher diagnosis performance.

* Corresponding author.

E-mail address: zhangxin_ieee@163.com (X. Zhang).

Nomenclature

AHU	Air Handling Unit	OAH	Outdoor air relative humidity (%RH)
HVAC	Heating, Ventilation, and Air conditioning	MAT	Mix air temperature (°C)
FDD	Fault detection and diagnosis	MAH	Mix air relative humidity (%RH)
ReLU	Rectified Linear Unit	SAT	Supply air temperature (°C)
PC	Pressure controller	SAH	Supply air relative humidity (%RH)
RF	Return fan	CWT	Chiller water temperature (°C)
SF	Supply fan	SAR	Supply airflow rate (m ³ /h)
OA	Outdoor Air	DAL	Duct air leakage (DAL)
RA	Return Air	FED	Fan efficiency decrease (FED)
SA	Supply Air	CVS	Cooling coil valve stuck (CVS)
EA	Exhausted Air	ODE	Outdoor damper excess (ODE)
OAT	Outdoor air temperature (°C)		

Recently, convolutional neural networks (CNNs) are commonly used to extract feature of time-series due to fewer parameters, high computation efficiency and classification performance. Motivated by the success of CNNs in various classification and recognition tasks, 1-dimensional convolutional neural networks (1DCNNs) had been used on raw sensor signal for bearing fault detection in [17], structural damage detection in [18], and modular multilevel converter (MMC) circuits fault diagnosis in [19]. Besides, the fault diagnosis of the power system and building sector were studied extensively [20–23]. These research works demonstrate that the 1DCNNs-based method can automatically extract discriminative feature representation of 1-dimensional signal for fault diagnosis, which reduces the dependence on hand-crafted feature extraction. However, these abovementioned methods are difficult to efficiently capture features of multiscale monitoring signals from the AHU system. For instance, for slow-changing temperature signal, the scale window of kernel should be large enough for the CNNs to capture the tendency of fault occurrence. Instead, for fast-changing airflow signal, the scale window of kernel should be small enough for the CNNs to detect sudden fault occurrence.

Inspired by the inception structure proposed by Szegedy [24,25], a new Multiscale CNNs (MCNNs) model is utilized in this paper to design an end-to-end multiscale feature learning structure for the fault diagnosis of AHU with multiscale monitoring signals. The proposed MCNNs-based FDD method can capture abundant discriminative feature by using multiscale kernels and improve the FDD performance for AHU system by the way of an end-to-end learning. The advantages of the work presented in this paper can be summarized as follows:

(a) A new end-to-end FDD method is proposed for the fault diagnosis of AHU system. The proposed method can work well with multiscale monitoring signals, which can be a general-purpose approach for intelligent FDD.

(b) The developed MCNNs model can automatically and effectively extract discriminative feature from the multiscale monitoring signals. This structure eliminates the requirement of complicated feature engineering in previous FDD methods.

(c) The proposed MCNNs method can improve the FDD performance by end-to-end learning. The end-to-end learning can alleviate the deterioration of performance caused by the separation of feature learning and classification and provide a promising way of FDD for the real-world AHU system.

The remainder of this paper is organized as follows: Section 2 describes the problem of the FDD for the AHU system. Section 3 proposes the MCNNs-based FDD method for the AHU system. Section 4 introduces the experiment platform of our lab and reports the experiment results. In the end, Section 5 summarizes the paper.

2. Problem description of FDD for AHU system

2.1. Operation principle of AHU

AHU is the most extensively operated component in the general HVAC system as Fig. 1(a). The mixed air, which is the mixture of the outdoor air (OA) and the return air (RA) from the buildings, is handled in AHU to supply air for buildings. As shown in Fig. 1(b), AHU includes three zones consisting of a mixed air zone, heat exchange zone and supply air zone. In the mixed air zone, the outside air (OA) and the return air (RA) are supplied to the AHU box respectively through the outside air inlet and return air inlet. The outside air damper and return air damper are operated simultaneously to control the proportion of the recirculated mixed air and maintain healthy and comfortable indoor air quality. The mixed air consists of OA and RA is handled by the filter to purify and equalize the air. In the heat exchange zone, the mixed air is cooled by cold exchange with the cooling coil connected to the chilled water tank, in which temperature is controlled at 10 °C. Analogously, in heating mode, the mixed air is heat up by heating coil. Then, the humidity of the exchanged mixed air is adjusted by the humidifier to meet the requirement of building environment. In the supply air zone, the supply fan adjusts motor speed to maintain the static pressure of supply air at its set-point with the variation of the loads. Many faults may be happened during this process due to equipment aging or damage.

2.2. Faults types and impacts of AHU system

AHU integrating many devices is prone to occur faults and causes a drop in efficiency. For example, duct air leakage causes reduced thermal conditioning capacity when the duct seal is incomplete. A decrease in the motor efficiency of the supply fan could reduce the supply airflow rate. The valve of the cooling coil stuck could reduce the efficiency of a temperature controller and waste more energy. The OA damper stuck could import excessive outdoor air and waste more energy in refrigeration mode. Taken together, Table 1 tabulates four common faults which result in the degradation of AHU system performance to test the performance of the proposed method.

Eight monitoring signals from sensors in our AHU experiment platform can be captured to diagnose the abovementioned faults. These signals includes: outdoor air temperature OAT (°C) and outdoor air relative humidity OAH (%RH), mixed air temperature MAT (°C) and mixed air relative humidity MAH (%RH), chiller water temperature CWT (°C), supply air temperature SAT (°C), supply air relative humidity SAH (%RH) and supply airflow rate SAR (m³/h). These signals with different change rates are shown in Fig. 2.

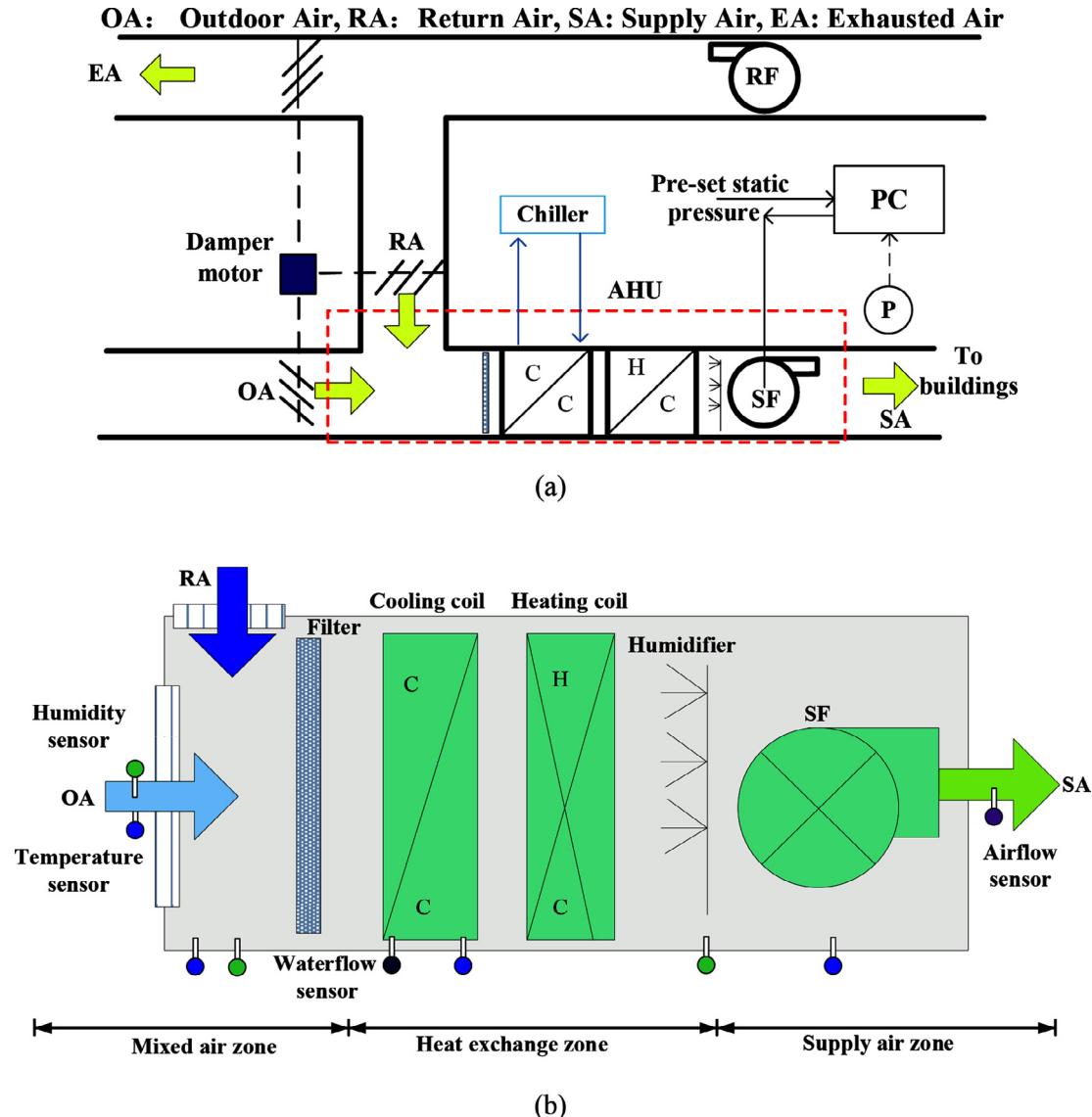


Fig. 1. Structure schematic of general AHU system: (a) general structure schematic of the HVAC system, (b) composition schematic of the AHU.

Table 1
Four faults occurring in the AHU system.

Fault number	Fault type	Mark symbol
1	Duct Air leakage	DAL
2	Fan efficiency decrease	FED
3	Cooling coil valve stuck	CVS
4	Outdoor damper excess	ODE

Table 2 summarized these eight signals, and it can be found that temperature and humidity signals are slow-changing signals, rather, airflow rate SAR is a fast-changing signal. Moreover, these signals have different amplitudes of variation. Due to the multiscale characteristic of these AHU signals, it is difficult to capture the comprehensive characteristic of fault only using a single-scale kernel. For example, the large kernel scale can capture the tendencies of slow-changing signals and a small kernel scale can capture local characteristics of fast-changing signal. It is a challenge to capture these different scale characteristics helpful for fault detection and diagnosis.

2.3. Challenges of FDD method for AHU system with multiscale monitoring signals

As summarized in **Table 3**, researchers have been investigating and proposing various solutions of fault diagnosis in HVAC systems to solve thermal discomfort and energy inefficiency in buildings. These previous methods have some disadvantages to deal with the FDD of AHU with multiscale monitoring signals. For instance, PCA is difficult to represent the abstract feature of multiscale signals due to the weakness of linear mapping. Wavelet transform (WT) is difficult to realize end-to-end learning with the classifier. Neural networks methods (NNs) are difficult to train the model for high diagnosis performance due to too many weight parameters. SVM and ELM has limited performance capability for multiscale monitoring signals since that single kernel function is not suitable for multiple different signal forms, and multiple kernel learning requires high computational cost. CNNs with relatively few parameters and high computation efficiency is suitable to extract feature for fault detection and diagnosis of AHU system, but it does not obtain a high accuracy due to the mismatching of

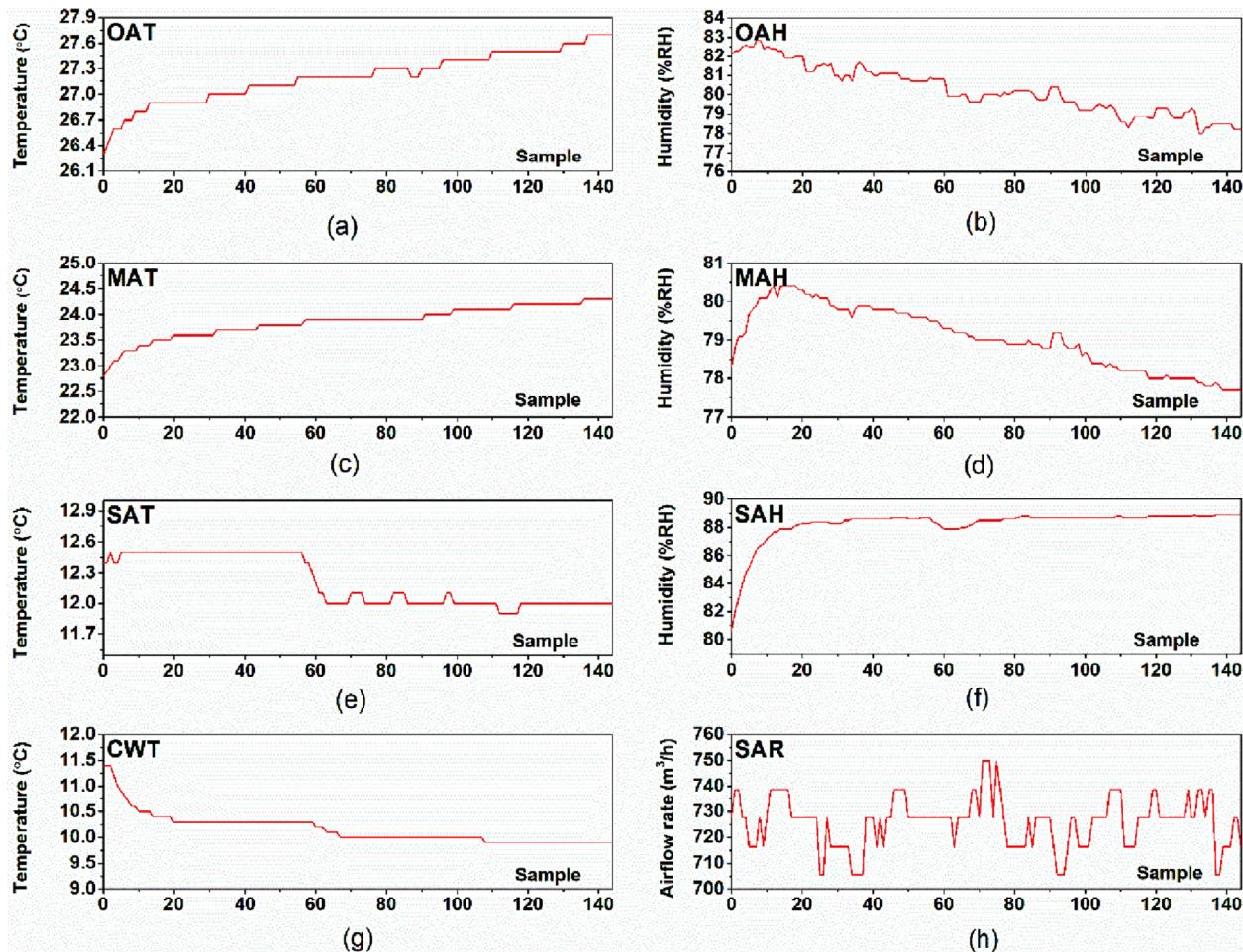


Fig. 2. Monitoring signals with different change rates on condition of 10 °C chiller water and 12 °C supply air: (a) OAT signal, (b) OAH signal, (c) MAT signal, (d) MAH signal, (e) SAT signal, (f) SAH signal, (g) CWT signal, (h) SAR signal. The horizontal axis represents the sample point with the sampling interval of 60 s, and the vertical axis represents the measurement value of signal.

Table 2
The description of 8 monitoring signals.

Name	Description	Amplitude of variation	Rate of variation
OAT	Outdoor air temperature	[26–28] °C	Slow
OAH	Outdoor air humidity	[78–82] %RH	Slow
MAT	Mix air temperature	[23–24.5] °C	Slow
MAH	Mix air humidity	[78–80] %RH	Slow
SAT	Supply air temperature	[12–12.5] °C	Slow
CWT	Chiller water temperature	[10–11.5] °C	Slow
SAH	Supply air humidity	[80–88] %RH	Slow
SAR	Supply airflow rate	[700–760] m³/h	Fast

Table 3
Summary of FDD methods for HVAC in the literatures.

Method	Limitation
Principal component analysis (PCA)	Linear mapping is difficult to represent the abstract feature of multiscale signals.
Neural networks (NNs)	Model has too many parameters and is difficult to train to obtain high diagnosis performance
Wavelet transform-NNs (WT-NNs)	It is difficult to realize end-to-end learning
SVM and ELM	Kernel function is difficult to apply to multiple different scale signals.
Convolution neural networks (CNNs)	Single-scale kernel does not match multiscale monitoring signals

single scale kernel and multiscale monitoring signals. Therefore, MCNNs-based fault detection and diagnosis method is proposed to solve this problem in the paper.

3. The proposed MCNNs-based FDD method for AHU system

3.1. Description of the proposed FDD method for AHU system

Flow chart of the FDD method for AHU system illustrated in Fig. 3(a), consists of an offline/online model training using the acquired data from AHU system, and an online diagnosis of real-time data with the trained model. The model parameters can be learned by offline/online training, then, the online real-time data are input into the offline/online trained model which will output the predicted result of fault type.

The proposed MCNNs-based FDD method for the AHU system is shown in Fig. 3(b). The data acquired from the database of our building management system are used to train/update the MCNNs model, and the trained model will be adopted to diagnose faults using multiscale monitoring signals. The proposed MCNNs-based FDD method is general and flexible, which may have different layer sizes and output conditions depending on the length of the input signals and the number of fault classes. In this paper, the proposed MCNNs-based FDD method aims to identify five AHU system conditions which include one normal condition and

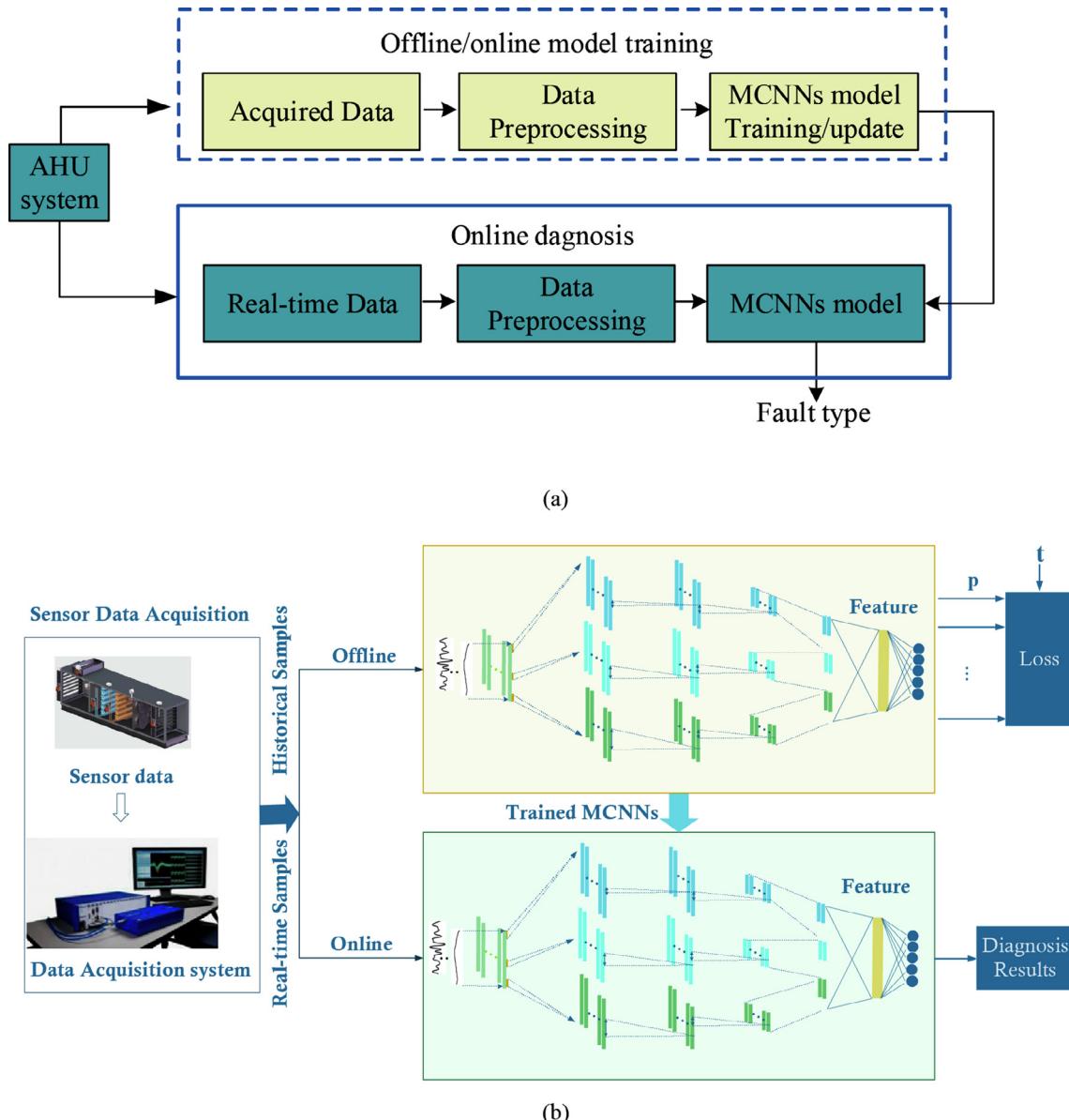


Fig. 3. Flow chart of the proposed MCNNs-based FDD method for the AHU system: (a) flow chart of the proposed FDD method, (b) structure chart of the proposed MCNNs-based FDD method.

four fault conditions: duct air leakage (DAL), fan efficiency decrease (FED), cooling coil valve stuck (CVS), and outdoor damper excess (ODE).

The MCNNs model is to automatically learn high-level discriminative feature helpful for the FDD of the AHU system directly from multiscale monitoring signals instead of manually designed feature extraction method such as PCA and wavelet transform [10,15]. The model trained by an end-to-end learning way can feedback the classification information to adjust the model parameters of feature learning. This way can improve the compatibility of feature learning and classification to extract better feature marked yellow in Fig. 3(b) for fault diagnosis. Therefore, the MCNNs model which can extract more discriminative feature to improve the diagnosis performance is a key component of the proposed MCNNs-based FDD method, and it will be introduced in detail as follows.

3.2. The proposed MCNNs model

3.2.1. Preliminary of MCNNs

Convolution has been a well-established approach for signal processing [26,27], and it is good at capturing characteristics of time series signal. Assume that K is a kernel of size s , and T is a time series signal of length N . Let $T * K$ denote the output result 1-Dimensional discrete convolution. Then, the i -th element of the output can be given as

$$(T * K)_i = \sum_{j=0}^{N-1} T_j K_{i-j} \quad (1)$$

The convolution can extract many insightful characteristics from the input time-series signal through the proper kernel. To automatically capture deeper abstract characteristics, convolu-

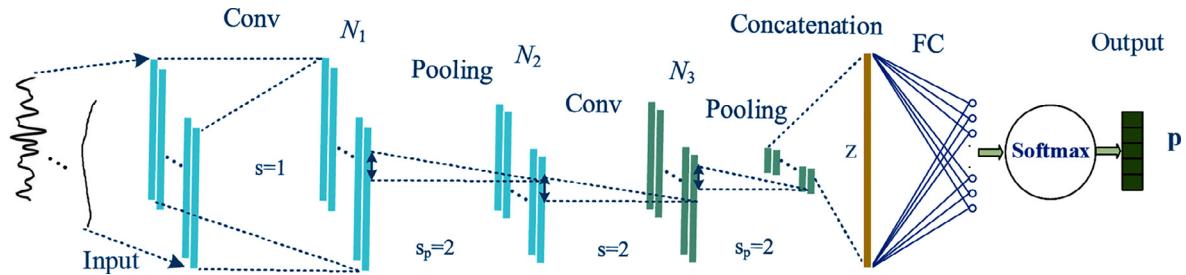


Fig. 4. 1D convolutional neural networks for 1D signal classification.

tional neural networks (CNNs) is designed by simulating the model for the mammalian visual cortex, which can extract advanced feature representation and output the probability of class by a feed-forward process as shown in Fig. 4. Generally, it consists of an input layer, the convolution layers (Conv), the pooling layers (Pooling), the fully connected layers (FC), and the Softmax layer. The parameters of CNNs are updated by the backpropagation algorithm [28] to minimize the loss of misclassification.

The convolution layer is the core block of the CNNs, and it can produce the output feature map by using the kernel sliding over the input feature map. Generally, a pooling layer is connected behind the convolution layer to reduce the spatial size of the output feature map for good robustness and smaller parameter size. In the pooling layer, subsampling, max-pooling or average-pooling can be adopted according to the actual need. The last part of the network is the fully connected layer to integrate the features and followed by the Softmax layer to output the probability of fault type as

$$\mathbf{p}_j = \frac{e^{\omega_j^T \mathbf{z}}}{\sum_{j=1}^C e^{\omega_j^T \mathbf{z}}} \quad (2)$$

where \mathbf{p} is the output probability vector of the fully connected Softmax layer with the input \mathbf{z} , \mathbf{p}_j is the j -th element of the output vector \mathbf{p} , ω_j is the weight parameter linked with the j -th output element and C is the number of output classes.

There are several hyper-parameters in the convolution layer, including the size of the kernel $s \in R$ which defines the scale of the convolution, the number of kernels $N_l \in R$ which is also the number of feature maps of the l layer, the padding and stride which

can be “same” or “valid” [29]. Besides, the type of activation function can be selected from Sigmoid or ReLU activation function, in which ReLU is commonly used with CNNs to avoid vanishing gradient problem [30].

The training of CNNs is adjusting kernel weights and hyper-parameters to minimize the cross-entropy loss given as

$$J = \sum_{i=1}^M \sum_{j=1}^C t_j^i \mathbf{p}_j^i \quad (3)$$

where t_j^i is the j -th element of the i -th sample target \mathbf{t}^i with one-hot encoding, \mathbf{p}_j^i is the j -th element of the i -th sample output \mathbf{p}^i and M is the number of samples. In the training process, other regularization techniques can be applied to avoid over-fitting such as batch normalization, and dropout technique.

3.2.2. Structure of the proposed MCNNs model

This paper focuses on the FDD of AHU on eight multiscale monitoring signals. In such cases, some fault states are difficult to be distinguished only simply extracting the fault feature with single-scale CNNs. To better capture more rich and discriminative characteristics of these monitoring signals for the improvement of FDD performance, a new MCNNs model is developed. The overall structure of the proposed MCNNs model is illustrated in Fig. 5, and it consists of three subparts as below:

(1) Input

The input of the proposed MCNNs model consists of 8 monitoring signals. Then, given a 8 by 8 matrix $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4, \mathbf{z}_5, \mathbf{z}_6, \mathbf{z}_7, \mathbf{z}_8]$ represents the 8 input data segments with a length of 8 sample

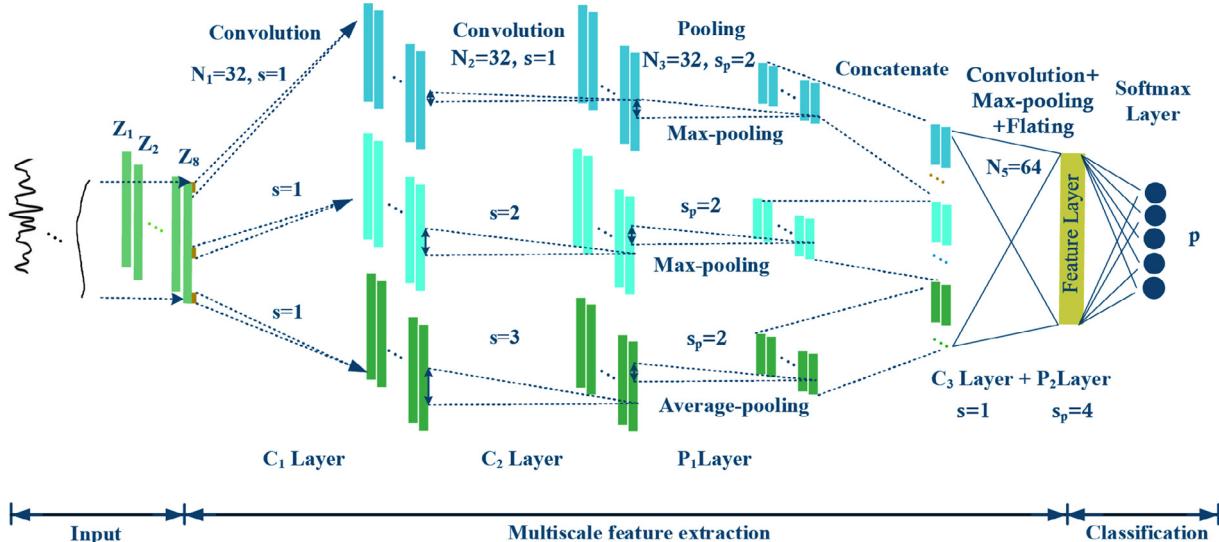


Fig. 5. Structure of the proposed MCNNs model.

points, where \mathbf{z}_1 represents the segment of OAT, \mathbf{z}_2 represents the segment of OAH, \mathbf{z}_3 represents the segment of MAT, \mathbf{z}_4 represents the segment of MAH, \mathbf{z}_5 represents the segment of CWT, \mathbf{z}_6 represents the segment of SAT, \mathbf{z}_7 represents the segment of SAH, and \mathbf{z}_8 represents the segment of SAR. Lots of data segments from acquired data will be input to the next convolution layer for feature extraction.

(2) Multiscale feature extraction

Because AHU monitoring signals are variable along with different seasons and weather conditions. The goal of MCNNs is to extract the robust discriminative features for fault diagnosis under different environments. The structure of MCNNs is designed to three parallel channels convolutional layers with multiscale kernels and following pooling layer.

As shown in Fig. 5, $\mathbf{Z} = [\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4, \mathbf{z}_5, \mathbf{z}_6, \mathbf{z}_7, \mathbf{z}_8]$ is input to three channels with C_1 and C_2 convolutional layers to extract discriminative feature and the P_1 pooling layer to improve the robustness of features. Then, the following C_3 convolutional layer and P_2 pooling layer are used to fuse the features from these three channels. The convolution kernel size of the C_1 Layer is set to 1 to transform the input to three channels and the weights in C_1 Layer can be trained to improve the performance of the following multiscale feature learning. The convolution kernel sizes of the C_2 Layer are set to 1, 2 and 3 from top to bottom to capture multiscale features from the feature maps of the C_1 Layer. These three different kernel sizes enable three-channel convolutional neural networks to learn changing feature on different scales respectively and these abundant features are very helpful for improving the classification of multi-scale signals.

The convolution kernel size of the C_3 Layer is set to 1 to fuse features from the previous three channels. These fused features have stronger distinguishing ability for the classification of multi-scale signals. In this way, these output feature maps can capture different scale features from the input signals. The whole process on the top branch is described in the following. The kernel slides with a window size of s over the whole input signals to extract local features. Accordingly, the i -th output feature map of the l layer is defined as

$$\mathbf{z}_i^l = \sigma(\mathbf{z}^{l-1} * W_i^l) = \sigma(\mathbf{a}_i^l) \quad (4)$$

where W_i^l represents the parameters of i -th kernel of the l layer, $*$ represents the operation of convolution, and $\sigma(\cdot)$ is a nonlinear activation function. In this study, ReLU ($\sigma(x) = \max(0, x)$) is adopted as the nonlinear activation function to prevent the problem of vanishing gradient. Using the operation as (4), the output feature maps of the first convolution layer can be obtained as $\mathbf{z}^1 = (\mathbf{z}_1^1, \mathbf{z}_2^1, \dots, \mathbf{z}_{N_1}^1)$ with input signal ($\mathbf{z}^0 = \mathbf{Z}$). Taking the output of the first convolution layer as the input of the second convolution layer, the output feature maps of the second convolution layer $\mathbf{z}^2 = (\mathbf{z}_1^2, \mathbf{z}_2^2, \dots, \mathbf{z}_{N_2}^2)$ can be obtained by repeating the same operation as the first convolution layer. Then, the pooling layer (P_1) is further applied to the feature maps \mathbf{z}^2 generated by two previous convolutional layers (C_1 and C_2), thus enabling to extract robust and location-invariant features. The max-pooling or average-pooling with a pooling length s_p is adopted for calculating local max value or mean value over the feature map. Then, the pooled feature map \mathbf{z}^3 can be obtained with the i -th element as

$$\mathbf{z}_i^3 = (\mathbf{z}_{i1}^3, \mathbf{z}_{i2}^3, \dots, \mathbf{z}_{i(N_2/s_p)}^3), i = 1, \dots, N_3$$

$$\mathbf{z}_{ij}^3 = \max_{(j-1)s_p+1 \leq k \leq js_p} \{\mathbf{z}_{ik}^2\} \text{ or } \mathbf{z}_{ij}^3 = \frac{1}{(j-1)s_p+1 \leq k \leq js_p} \sum_{k=1}^{(j-1)s_p+1 \leq k \leq js_p} \{\mathbf{z}_{ik}^2\}.$$

After two convolutional layers (C_1 and C_2) and pooling layer (P_1), we simply concatenate these obtained feature maps from each branch to get a concatenated aggregation feature, which is denoted

as $\mathbf{z}^4 = (\tilde{\mathbf{z}}^3; \tilde{\mathbf{z}}^3; \mathbf{z}^3)$, where $\tilde{\mathbf{z}}^3$ represents the feature maps from the top branch, $\tilde{\mathbf{z}}^3$ represents the feature maps from the middle branch,

where \mathbf{z}^3 represents the feature maps from the bottom branch. The concatenated aggregation feature is input into the following layer of Convolution + Max-pooling + Flating including a convolution layer with 64 output feature maps and $s = 1$, a max-pooling layer with $s_p = 4$ and a flat operation. Finally, a syncretic feature vector $\bar{\mathbf{z}}$ with size of 64 can be obtained as the multiscale feature representation. Compared to single-scale features extracted from the input signals, multiscale features may contain more complementary and discriminative fault pattern features, and these complementarity features are helpful for the classification of multiscale monitoring signals.

(3) Classification

The AHU fault diagnosis studied in this paper can be regarded as a multiclass classification problem. The multiscale feature representation $\bar{\mathbf{z}}$ obtained in the multiscale feature extraction stage is directly fed to the next fully connected Softmax layers consisting of a fully connected layer with the ReLU activation function and the following Softmax function, which outputs a class conditional probability \mathbf{p} of the input. Assuming that there are C classes of AHU conditions for the syncretic feature representation $\bar{\mathbf{z}}$, corresponding output probability $p_i \in [0, 1]$ of class i can be calculated by

$$\mathbf{p}_i = \frac{e^{(\sigma(W_i^T \bar{\mathbf{z}}))}}{\sum_{i=1}^C e^{(\sigma(W_i^T \bar{\mathbf{z}}))}} \quad (5)$$

where W_i is the model parameter needed to be learned in a fully connected layer and the sum of the output probabilities is 1 ($\sum_{i=1}^C p_i = 1$).

3.2.3. Training of the MCNNs model

The MCNNs model is trained using the backpropagation (BP) [31] and Adam optimization algorithm [32]. The cross-entropy between the predicted labels and the true target labels is adopted as the loss function J . The backpropagation expression of convolution layer can be deduced as

$$\delta^{l-1} = \frac{\partial J}{\partial \mathbf{a}^{l-1}} = \delta^l \cdot \frac{\partial \mathbf{a}^l}{\partial \mathbf{a}^{l-1}} = \delta^l * \text{rot180}(W^l) \odot \sigma'(\mathbf{a}^{l-1}) \quad (6)$$

where $\mathbf{a}^l = \sigma(\mathbf{a}^{l-1}) * W^l = \mathbf{z}^{l-1} * W^l$, δ^{l-1} is the loss gradient of the $l-1$ layer, δ^l is the loss gradient of the l layer, $\text{rot180}(\cdot)$ represents the kernel rotating 180 degrees, \odot represents element-wise product. Further, the gradient of model weights can be obtained as

$$\frac{\partial J}{\partial W^l} = \frac{\partial J}{\partial \mathbf{a}^l} \cdot \frac{\partial \mathbf{a}^l}{\partial W^l} = \delta^l * \text{rot180}(\mathbf{z}^{l-1}) \quad (7)$$

Similarly, the backpropagation equation for pooling layer can be deduced as

$$\delta^{l-1} = \frac{\partial J}{\partial \mathbf{a}^{l-1}} = \frac{\partial J}{\partial \mathbf{z}^{l-1}} \cdot \frac{\partial \mathbf{z}^{l-1}}{\partial \mathbf{a}^{l-1}} = \text{upsample}(\delta^l) \cdot \sigma'(\mathbf{a}^{l-1}) \quad (8)$$

where $\text{upsample}(\delta^l)$ is an upsampling operation, \cdot represents the operation of restoring the original maximum position for the max-pooling or the operation of average for the average-pooling as shown in Fig. 6. The backpropagation of the last fully connected layer is similar with the conventional neural networks [28], then, all the backpropagation methods can be obtained. To minimize the loss function with high computation and little memory, Adam optimization algorithm is employed to update model parameters by the obtained gradient of model weights. Due to the model is trained by Adam which is an improved version of stochastic gradient

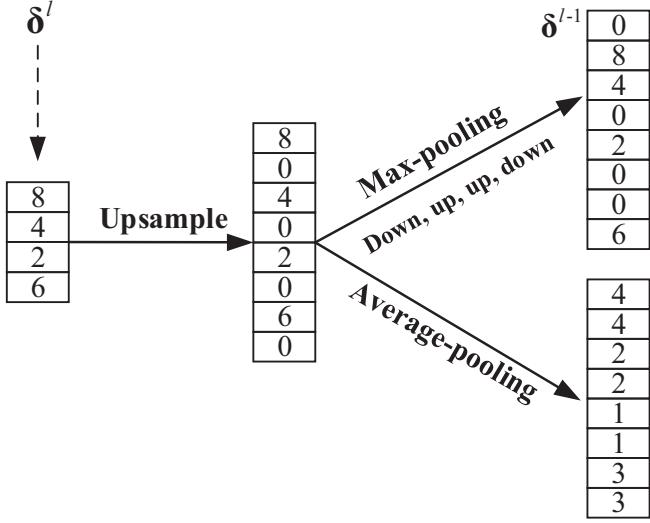


Fig. 6. An example of the backpropagation method for pooling layer.

descent (SGD) method, therefore, online learning can be enabled. The model parameters can be updated using the new acquisition data and their states.

4. Experiment results on our AHU system

In this section, the proposed MCNNs-based FDD method is evaluated on an AHU experiment platform in our lab as shown in Fig. 7. The description of data acquisition (DAQ) is first introduced, and then the model is trained using a laptop with 32 GB RAM and pro-

cessor Intel(R) Core (TM) i7-7700HQ CPU at 2.8 GHz speed. Finally, the detailed experiment results of FDD are shown and discussed.

4.1. Data acquisition

AHU usually encounters faults during actual operation, and four common faults are simulated on our AHU system. In the simulation on our AHU platform, every fault is simulated under three magnitudes, for example, DAL is simulated with the opening sizes on the side of the AHU box are 1 cm, 2 cm and 3 cm respectively. FED is simulated by turning off 1, 2, and 3 fans respectively to simulate three different magnitudes. CVS and ODE are simulated by adjusting three different angles of the control valve. The experiment dataset was acquired under the normal state and four simulation fault states on many different environments and filtered to remove some outliers. The description of acquisition dataset is listed in Table 4.

The total acquisition dataset includes 6048 data segment samples, in which 1324 samples are acquired from normal operating condition, and the sample sizes of the other four faults are respectively 1580, 1324, 852, 968 as listed in Table 4. The violin figures [33] of these monitoring data are displayed in Fig. 8. It can be found that these data distributions are irregular because of the change of environment and magnitude, and it is difficult to distinguish different faults only using the value of one signal data. Complementary information in these data are required to fuse to distinguish different fault scenarios. Therefore, MCNNs model is necessary to be proposed to fuse these different scale features for the FDD of AHU. The total dataset is split into two parts, and one part of them, including 2432 samples, is treated as training and validation data, and the other part, including 3616 samples, is treated as testing data.



Fig. 7. AHU test platform: (a) front of the AHU, (b) back of the AHU.

Table 4

Description of acquisition dataset.

Label	Fault type	Fault description	Fault position	Sample size
1	Normal	Normal condition	No	1324
2	DAL	Duct air leakage	Duct surface	1580
3	FED	Fan efficiency decrease	Fan circuits	1324
4	CVS	Cooling coil valve stuck	Cooling coil	852
5	ODE	Outdoor air excess	Damper	968

4.2. Results of loss curve and feature learning ability

In this paper, three scales are considered in the proposed MCNNs model to extract abundant discriminative features helpful for fault classification. The input data including 8 data segments are fed into the MCNNs model. The number of kernels for all the convolutional layers (C_1 , C_2 , and C_3) are set to 32, and the pooling lengths for pooling layers (P_1 and P_2) are respectively set to 2 and 4. The feature representation with size of 64 is output by the input data going through all the convolutional and pooling layers. In the last stage of classification, the output feature is fully connected to the final output layer with size of 5, corresponding to the number of considered running conditions of the AHU system. The proposed MCNNs model is trained by Adam gradient descent algorithm with a batch size of 100 to minimize the loss function. In the training process, the learning rate is initialized to 0.001, and the training epoch is set to 30. NNs, WT-NNs and MCNNs are trained to minimize the loss function using Adam optimizer and loss curves of these three neural network models are presented in Fig. 9.

As demonstrated in the Fig. 9, the horizontal axis represents the training epoch, and the vertical axis represents the average loss value in the corresponding epoch. At the beginning of training, the average loss value drops greatly, indicating that the learning rate is appropriate, and the gradient descent process is performed. After a certain stage of learning, the loss curve tends to be stable. Compared with two curves of Fig. 9(a) and Fig. 9(b), the loss curve of MCNNs is the smoothest and the loss value of MCNNs is also the lowest as shown in Fig. 9(c). The comparison of these loss curves shows that the MCNNs model has the best loss function value and best stability.

To better observe the benefits of multiscale features extracted by the proposed MCNNs model, the dimension of the learned feature representation is reduced to two dimensions using the t-SNE technique [34] for visualization. The 2D feature representations of the validation data are shown in Fig. 9, where 2D feature points represent different states described by different colors. By comparing Fig. 10(a) and Fig. 10(b), it can be found that using the proposed MCNNs model can better separate features from different types which are helpful for classification and imply strong generalization performance of the proposed model.

4.3. Comparison results of diagnosis performance

4.3.1. Performance metrics: confusion matrix and F1-score

The fault diagnosis issue studied in this paper corresponds to a problem of five-class classification (five states in our case). To compare the performance of classification, two metrics are adopted to validate the performances of different methods.

The first is the confusion matrix which is a useful tool to show the corresponding results of the true labels and predicted labels using different methods. The row of confusion matrix represents the true label i ($i \in \{1, 2, 3, 4, 5\}$) and the column represents the predicted label j ($j \in \{1, 2, 3, 4, 5\}$), the corresponding matrix element is the number of samples with label i predicted to the label j .

The second is F1-score [35] which commonly used to numerically measure the overall binary classification performance is adopted as the metric of diagnosis performance of each fault for different methods. F1-score is defined as

$$F1 - score = \frac{2TP}{2TP + FP + FN} \quad (9)$$

where TP (True Positive) is the number of exactly classified as positive samples, FP (False Positive) is the number of misclassified as positive samples, TN (True Negative) is the number of exactly classified as negative samples, and FN (False Negative) is the number of misclassified as negative samples. In the following studies, the F1-score for each fault is respectively calculated according to the binary classification method, and the minimum results will be reported to compare.

4.3.2. Comparison results with existing methods

The proposed method is compared with other five commonly used methods which are NNs, PCA + SVM, SVM, WT-NNs, ELM, respectively. Neural Networks (NNs) model is designed with six layers in which the number of neurons is 64, 64, 32, 16, 8, 5, respectively. In PCA + SVM model, the number of selected principal components is set to retain 90% of the overall variances and the kernel parameter is selected from the interval of [10–2, 10–1, 1, 10, 100] using the maximum selection method on validation data. In WT-NNs model, the layer number of wavelet decomposition is 3 and the NNs structure is the same with the above NNs model. In ELM model, the kernel parameter is selected from the interval of [10–2, 10–1, 1, 10, 100] using the maximum selection method on validation data.

Confusion matrixes of our comparison experiments are shown in Fig. 11. In the upper left subfigure Fig. 11(a), the first row shows that 530 segments of Normal condition are correctly classified, and 6 segments of Normal condition are misclassified to CVS fault. The second row shows that all 848 segments of DAL fault are correctly classified. The third row shows that 1 segment of FED fault is misclassified to Normal condition, and 551 segments of FED fault are correctly classified. The fourth row shows that 5 segments of CVS fault are misclassified to Normal condition, and 899 segments of CVS fault are correctly classified. The last row shows that all 776 segments of ODE fault are correctly classified. Classification performance in other subfigures can be similarly analyzed and summarized. The final synthesis of all figures is that the confusion matrix of the MCNNs method has the largest diagonal element values compared with that of other methods, and this implies that the proposed MCNNs method can extract strong discriminative features and obtain higher diagnosis performance compared with other methods.

The comparison results among the MCNNs method and other five methods are listed in Table 5, in which the leftmost column records the test time, and the other columns record the test score of different methods. The test time is listed to compare the computational complexity of the mentioned methods. For kernel-based methods, SVM, PCA + SVM and ELM, the computation complexities have a quadratic relationship with the number of samples ($O(M^2)$, M is the number of samples). For NN-based methods, MCNNs, NNs,

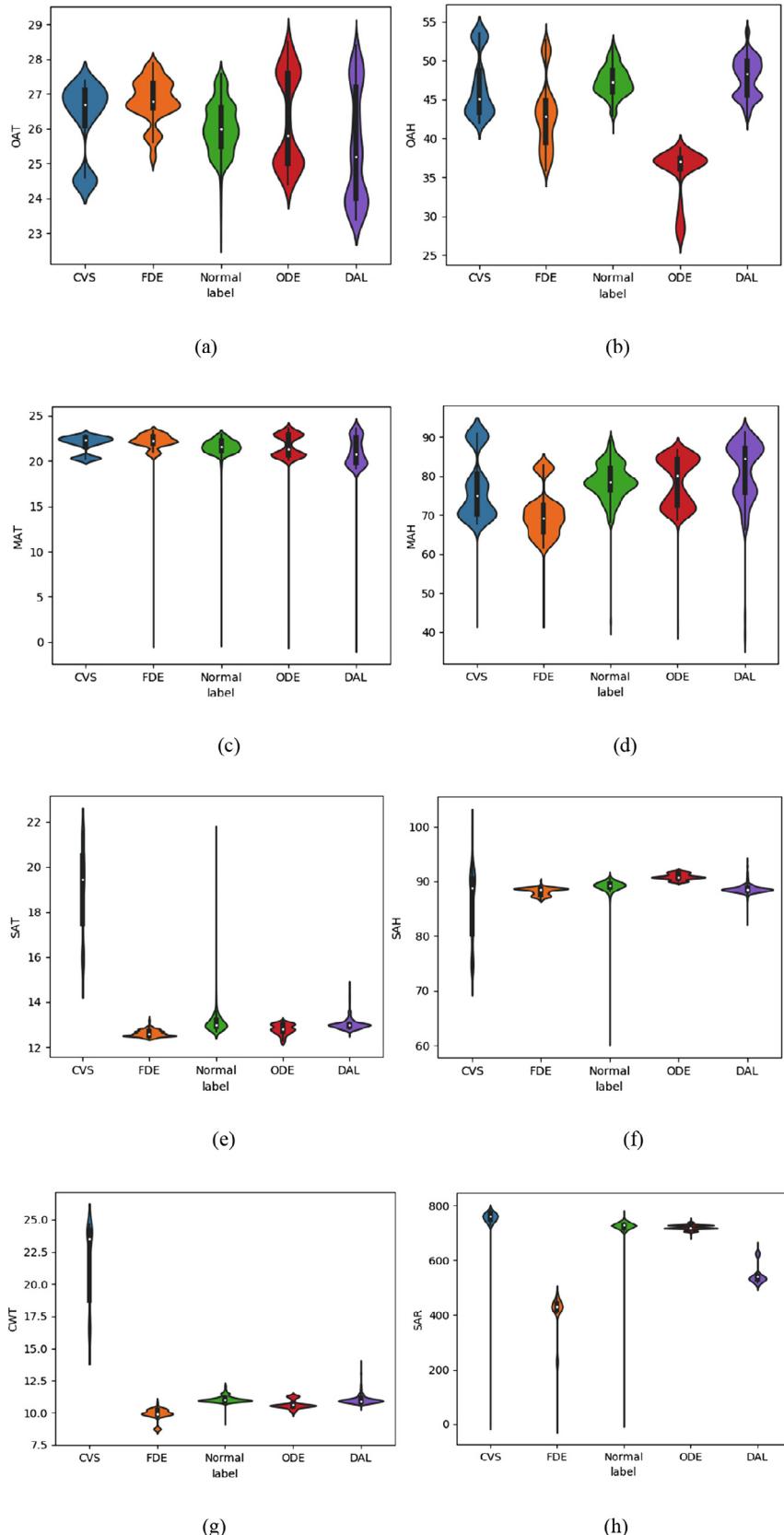


Fig. 8. The violin figures of the acquisition dataset: (a) the violin figure of OAT, (b) the violin figure of OAH, (c) the violin figure of MAT, (d) the violin figure of MAH, (e) the violin figure of SAT, (f) the violin figure of SAH, (g) the violin figure of CWT, (h) the violin figure of SAR. The violin figure describes the data distribution on each fault, in which the horizontal axis represents the label of the fault, and the vertical axis represents the value of the signal data.

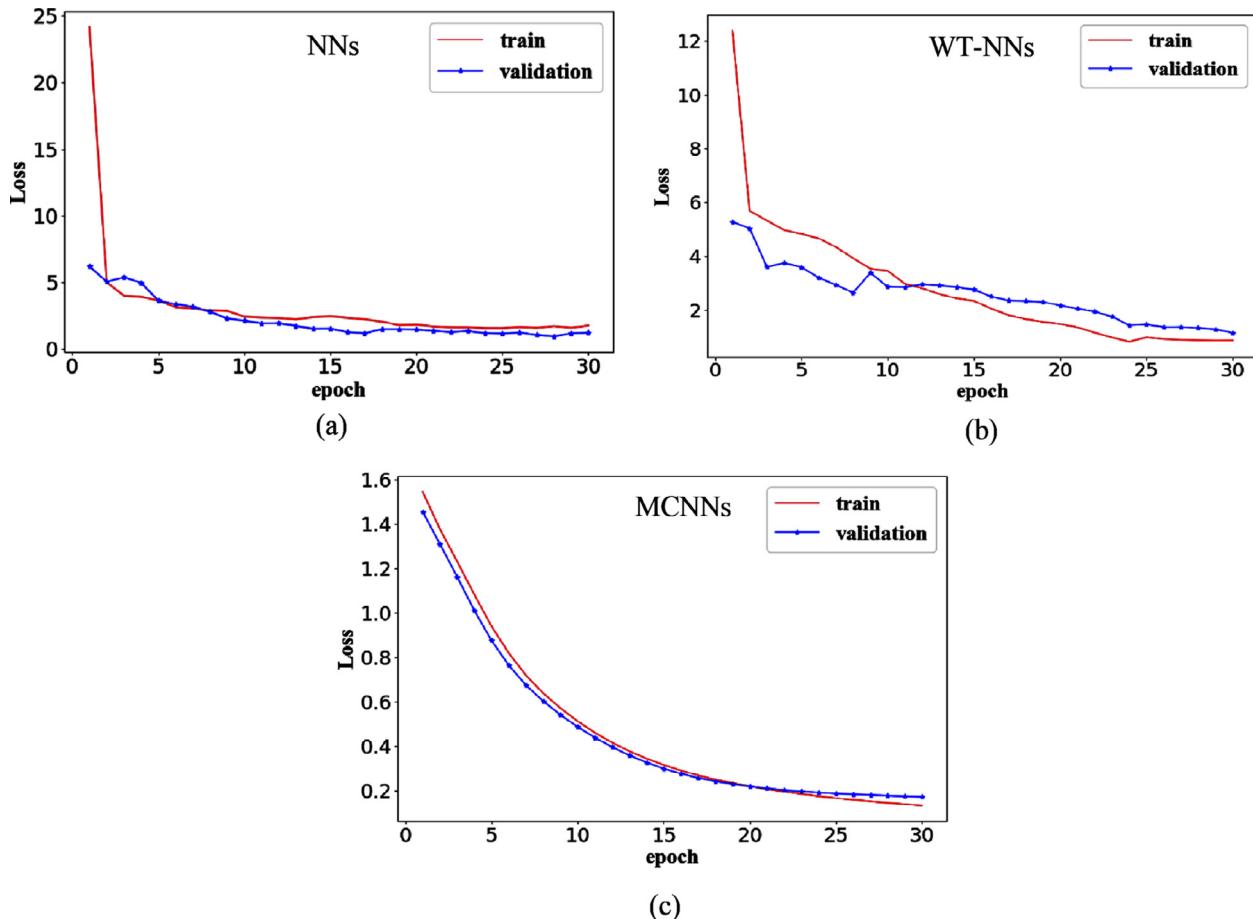


Fig. 9. Loss curves of three neural networks models: (a) NNs model, (b) WT-NNs model, (c) MCNNs model.

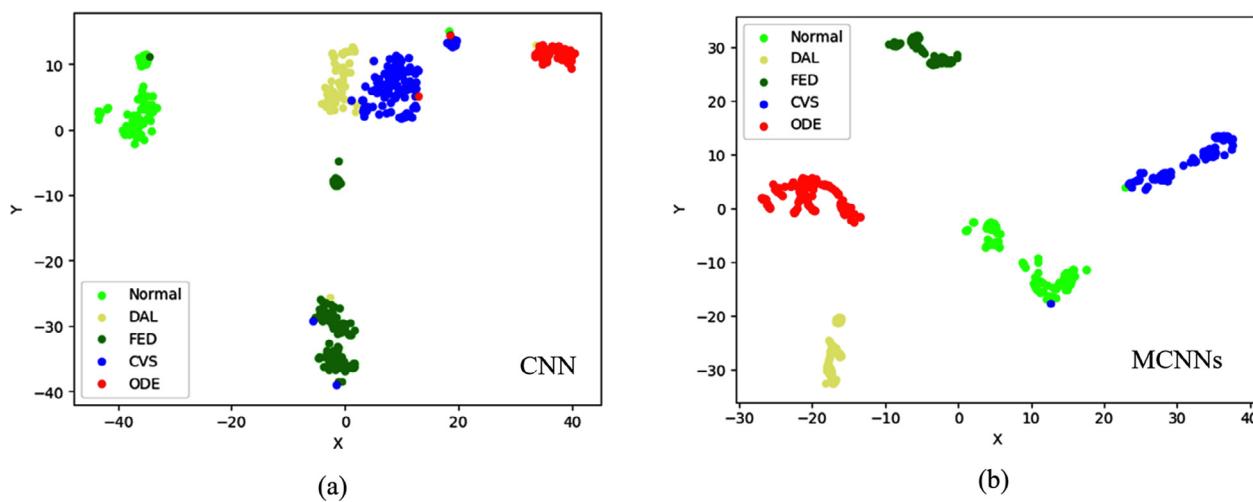


Fig. 10. 2D visualization of the extracted features using the t-SNE technique: (a) using CNN model in Fig. 4, (b) using the proposed MCNNs model. 2D points represent the sample features, and colors represent fault states.

WT-NNs, the computation complexities have a linear relationship with the number of samples ($o(M)$). These indicate that NN-based method can be extended for large scale dataset and kernel-based method is suited for relative small-scale dataset. For three NN-based methods, their computational complexities are proportional to the number of multiplications. Tab.5 shows that the MCNNs model takes 0.0079(s), the NNs model takes

0.0058(s), the PCA + SVM model takes 0.0372(s), the SVM model takes 0.0316(s), the WT-NNs model takes 0.0095(s) and the ELM model takes 0.0266(s) to detect all the test data. These illustrate that the test times of NN-based methods (MCNNs, NNs and WT-NNs) are comparable and kernel-based methods take more time to detect fault than NN-based methods.

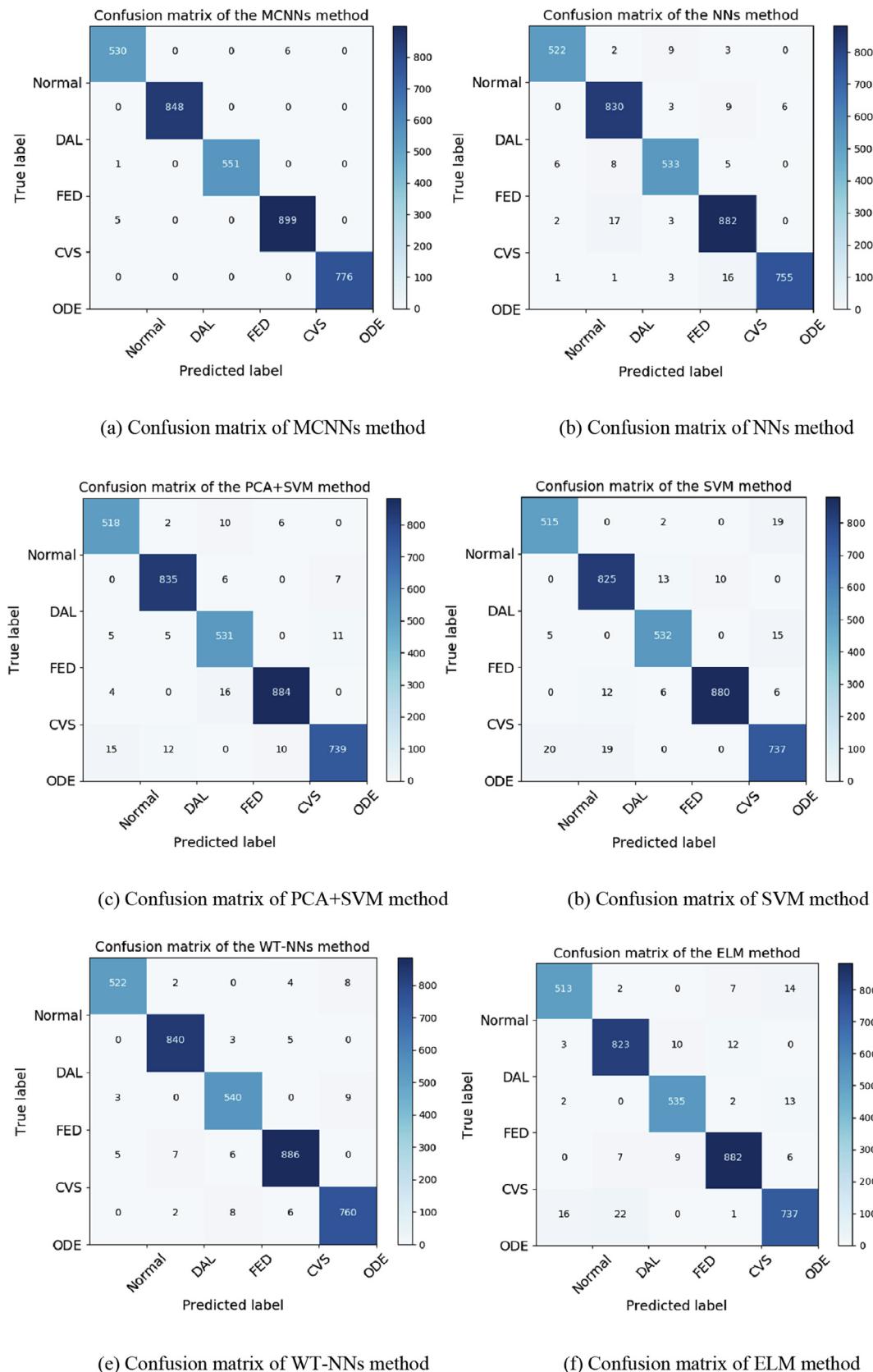


Fig. 11. Confusion matrixes of six methods: (a) MCNNs method, (b) NNs method, (c) PCA + SVM method, (d) SVM method, (e) WT-NNs method, (d) ELM method.

In terms of test scores, by dividing $(2 * \text{TP} = 1060)$ by $(2 * \text{TP} + \text{FP} + \text{FN} = 1060 + (1 + 5) + 6 = 1072)$, the F1-score of the MCNNs

method under Normal state is 0.989. It can be easily found that the F1-score 0.953 of ODE fault is the lowest value for the ELM

Table 5

Comparison results of F1-score metric (bold is the minimum of each method).

Test time (s)	Method\Condition	Normal	DAL	FED	CVS	ODE	Min. F1
0.0079	MCNNs	0.989	1	0.999	0.994	1	0.989
0.0058	NNs	0.978	0.973	0.966	0.970	0.982	0.966
0.0372	PCA + SVM	0.961	0.981	0.952	0.980	0.964	0.952
0.0316	SVM	0.957	0.968	0.963	0.981	0.949	0.949
0.0095	WT-NNs	0.979	0.989	0.974	0.982	0.979	0.974
0.0266	ELM	0.959	0.967	0.967	0.976	0.953	0.953

method, the F1-score 0.974 of FED fault is the lowest value for the WT-NNs method, the F1-score 0.949 of ODE fault is the lowest value for the SVM method, the F1-score 0.952 of FED fault is the lowest value for the PCA + SVM method, the F1-score 0.966 of CVS fault is the lowest value for the NNs method. In summary, for all kinds of faults, the proposed MCNNs method achieves the best overall performance higher than 0.989. This can be explained that the proposed multiscale fault diagnosis method can capture more discriminative feature by multiscale kernels. Moreover, in MCNNs method, the classification information can be fed back into the feature learning by the way of an end-to-end learning, which PCA + SVM, SVM, WT-NNs and ELM do not have. In other words, the independence of feature learning and classification may be lost useful discrimination information and the way of end-to-end learning can effectively capture more discriminative features for fault diagnosis.

5. Conclusion

This paper proposes a new MCNNs method of intelligent fault diagnosis for the AHU system with multiscale monitoring signals. The proposed method can perform end-to-end feature extraction and classification simultaneously without the need for complex handcrafted feature engineering. The diagnosis ability is improved by strong discriminative multiscale features and less classification information loss of end-to-end learning. Experiment results demonstrate that the proposed MCNNs method outperforms previous FDD methods for AHU system. The proposed FDD method can improve the reliability and energy efficiency of the AHU system. Besides, the proposed method offers a new general-purpose method for the field of fault diagnosis and can be easily extended to different machines and industrial systems.

In our future work, the scalability of the proposed MCNNs method on imbalanced data distribution will be verified. The imbalanced multiscale feature learning and FDD method will be investigated to mitigate the impact of skewed data distribution between normal and fault data, so that the general performance of the proposed method can be further improved.

Acknowledgements

This work is supported by NSFC (61976005, 61572032, 61902167, 61871204), Natural Science Foundation of Anhui Province (1908085MF215, 1908085QE247), Research Foundation of Education Department of Anhui Province (KJ2019A0149), Open Research Fund of AnHui Key Laboratory of Detection Technology and Energy Saving Devices (DTESD2020A03), and Open Research Fund of AnHui Polytechnic University (Xjky02201903).

CRediT authorship contribution statement

Fanyong Cheng: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft. **Wenjian Cai:** Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Xin**

Zhang: Writing - review & editing, Supervision, Visualization. **Huanyue Liao:** Writing - review & editing, Visualization. **Can Cui:** Writing - review & editing, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.enbuild.2021.110795>.

References

- [1] S.E. Kwan, R. Shaughnessy, U. Haverinen-Shaughnessy, et al., The impact of ventilation rate on the fungal and bacterial ecology of home indoor air, *Build. Environ.* (2020), 106800.
- [2] Y. Jiang, Z. Luo, Z. Wang, et al., Review of thermal comfort infused with the latest big data and modeling progresses in public health, *Build. Environ.* (2019), 106336.
- [3] "Energy technology perspectives 2017." Available: <https://www.iea.org/etp2017>.
- [4] Z. Shi, W. O'Brien, Development and implementation of automated fault detection and diagnostics for building systems: a review, *Autom. Constr.* 104 (2019) 215–229.
- [5] S. Deshmukh, S. Samouhos, L. Glicksman, et al., Fault detection in commercial building VAV AHU: a case study of an academic building, *Energy Build.* 201 (2019) 163–173.
- [6] M.S. Piscitelli, D.M. Mazzarelli, A. Capozzoli, Enhancing operational performance of AHUs through an advanced fault detection and diagnosis process based on temporal association and decision rules, *Energy Build.* (2020), 110369.
- [7] R. Yan, Z. Ma, G. Kokogiannakis, et al., A sensor fault detection strategy for air handling units using cluster analysis, *Autom. Constr.* 70 (2016) 77–88.
- [8] S. Wang, F. Xiao, AHU sensor fault diagnosis using principal component analysis method, *Energy Build.* 36 (2) (2004) 147–160.
- [9] S. Wang, F. Xiao, Detection and diagnosis of AHU sensor faults using principal component analysis method, *Energy Convers. Manage.* 45 (17) (2004) 2667–2686.
- [10] S. Li, J. Wen, A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform, *Energy Build.* 68 (2014) 63–71.
- [11] K. Yan, Z. Ji, H. Lu, et al., Fast and accurate classification of time series data using extended ELM: application in fault diagnosis of air handling units, *IEEE Trans. Syst., Man, Cybern: Syst.* 49 (7) (2017) 1349–1356.
- [12] J. Liang, R. Du, Model-based fault detection and diagnosis of HVAC systems using support vector machine method, *Int. J. Refrig.* 30 (6) (2007) 1104–1114.
- [13] F. Magoulès, H. Zhao, D. Elizondo, Development of an RDP neural network for building energy consumption fault detection and diagnosis, *Energy Build.* 62 (2013) 133–138.
- [14] Z. Du, B. Fan, X. Jin, J. Chi, Fault detection and diagnosis for buildings and HVAC systems using combined neural networks and subtractive clustering analysis, *Build. Environ.* 73 (2014) 1–11.
- [15] B. Fan, Z. Du, X. Jin, et al., A hybrid FDD strategy for local system of AHU based on artificial neural network and wavelet analysis, *Build. Environ.* 45 (12) (2010) 2698–2708.
- [16] K.-P. Lee, W.u. Bo-Huei, S.-L. Peng, Deep-learning-based fault detection and diagnosis of air-handling units, *Build. Environ.* 157 (2019) 24–33.
- [17] T. Ince, S. Kiranyaz, L. Eren, M. Askar, M. Gabbouj, Real-time motor fault detection by 1-D convolutional neural networks, *IEEE Trans. Ind. Electron.* 63 (11) (2016) 7067–7075.
- [18] O. Abdeljaber, O. Avci, M.S. Kiranyaz, B. Boashash, H. Sodano, D.J. Inman, 1-D CNNs for structural damage detection: verification on a structural health monitoring benchmark data, *Neurocomputing* 275 (2018) 1308–1317.

- [19] S. Kiranyaz, A. Gastli, L. Ben-Brahim, N. Alemadi, M. Gabbouj, Real-time fault detection and identification for MMC using 1D convolutional neural networks, *IEEE Trans. Ind. Electr.* (2018), 1-1.
- [20] A. Gupta, G. Gurrala, P.S. Sastry, An online power system stability monitoring system using convolutional neural networks, *IEEE Trans. Power Syst.* (2018), 1-1.
- [21] S. Hamdaoui, M. Hamdaoui, A. Allouhi, R.E. Alaiji, T. Kousksou, A.E. Bouardi, Energy demand and environmental impact of various construction scenarios of an office building in morocco, *J. Cleaner Prod.* 188 (2018) 113–124.
- [22] R. Yan, Z. Ma, G. Kokogiannakis, Y. Zhao, A sensor fault detection strategy for air handling units using cluster analysis, *Autom. Constr.* 70 (2016) 77–88.
- [23] Z. Du, B. Fan, J. Chi, X. Jin, Sensor fault detection and its efficiency analysis in air handling unit using the combined neural networks, *Energy Build.* 72 (2014) 157–166.
- [24] Szegedy, Christian, et al. Inception-v4, inception-resnet and the impact of residual connections on learning. Thirty-First AAAI Conference on Artificial Intelligence; 2017.
- [25] C. Szegedy et al., Rethinking the inception architecture for computer vision, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016.
- [26] S. Mallat, *A Wavelet Tour of Signal Processing*, Elsevier, 1999.
- [27] B. Shi, X. Bai, C. Yao, An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (11) (2016) 2298–2304.
- [28] D.E. Rumelhart, R. Durbin, R. Golden, et al., Backpropagation: the basic theory, *Backpropagation: Theory, architectures and applications* (1995) 1–34.
- [29] W. Wang, S. Dou, Z. Jiang, et al., A fast dense spectral-spatial convolution network framework for hyperspectral images classification, *Remote Sensing* 10 (7) (2018) 1068.
- [30] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines[C]. *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010: 807–814.
- [31] Z. Zhang, *Derivation of Backpropagation in Convolutional Neural Network (CNN)*, University of Tennessee, Knoxville, TN, 2016.
- [32] D.P. Kingma, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
- [33] J.L. Hintze, R.D. Nelson, Violin plots: a box plot-density trace synergism, *Am. Statist.* 52 (2) (1998) 181–184.
- [34] L. Maaten, G. Hinton, Visualizing data using t-SNE, *J. Machine Learn. Res.* 9 (Nov) (2008) 2579–2605.
- [35] H. He, E.A. Garcia, Learning from imbalanced data, *IEEE Trans. Knowl. Data Eng.* 21 (9) (2009) 1263–1284.