IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 20, 2021, accepted July 28, 2021, date of publication July 30, 2021, date of current version August 6, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3101471

Deep Flexible Transmitter Networks for Non-Intrusive Load Monitoring of Power Distribution Networks

CHENXIAO MA AND LINFEI YIN
College of Electrical Engineering, Guangxi University, Nanning 530004, China
Corresponding author: Linfei Yin (yinlinfei@163.com)

The work of Linfei Yin was supported by the Guangxi Natural Science Foundation under Grant AD19245001 and Grant 2020GXNSFBA159025.

ABSTRACT With the massive development of various new energy sources, the balance of supply and demand in the power grid faces a considerable challenge. A reliable way is to perform demand-side management on energies. For demand-side management, monitoring the load connected to power distribution networks is necessary. This paper proposes deep flexible transmitter networks to quickly monitor the load when the load is connected to power distribution networks. The proposed algorithm combines flexible transmitter networks and deep backpropagation neural networks. After measuring the waveforms of loads, the algorithm is trained by the measured operational data. The testing results show that the proposed deep flexible transmitter networks can accurately monitor the load connected to power distribution networks. Compared with the deep backpropagation neural networks, the proposed algorithm improves the monitoring accuracy by more than 5%. The speed of the proposed algorithm is verified in experiments. After testing on embedded devices, the proposed algorithm can satisfy the requirements of edge computing systems.

INDEX TERMS Non-intrusive load monitoring, deep backpropagation neural networks, flexible transmitter networks, load modeling, embedded calculation.

I. INTRODUCTION

With the increasing economic and social development, technological progress, and living standards, numerous kinds of loads have entered human life. The working principles and modes of these loads are different. Under the excitation of the power system voltage, loads have different responses. The massive development of electronic equipment has brought various harmonics to power distribution networks. Various harmonics cause the current waveform of power distribution networks to deviate from the sinusoidal waveform [1]. The connection of high-power loads causes voltage drops and frequency fluctuations in power distribution networks. Loads with electronic equipment or high-power hurt the safe operation of power distribution networks [2].

With the decline of renewable energy prices and the continuous improvement of pursuit, grid companies contribute to environmental protection [3], various renewable energies have been connected to power distribution networks [4]. One way to solve distributed power consumption problems and balance the power flow between the power supply and demand is demand-side management by dispatching the supply side [5] and anomalous behavior identification [6]. The load monitoring of power distribution networks is considered in the paper.

Load monitoring has two types, i.e., intrusive and non-intrusive types. Non-intrusive load monitoring is applied without modifying the load and installing sensor equipment on the access units or houses. In recent years, thanks to the improvement of computer technology and the development of intelligent algorithms [7], non-intrusive load monitoring had diversified choices and realized higher monitoring accuracy [8]. Liu *et al.* applied the template's waveform matching to monitor load [9]. Time-series data of power changes have been utilized for non-invasive load monitoring [10]. A post-processing method has been designed to improve the monitoring accuracy of existing non-intrusive load monitoring algorithms [11]. However, the robustness of the post-processing algorithm is not strong enough for other loads. The post-processing method has been combined with

The associate editor coordinating the review of this manuscript and approving it for publication was Mehdi Hosseinzadeh.

deep neural networks to improve the robustness of the monitoring algorithm [12]. Reference [13] designed a combined Bayesian method to improve the robustness of monitoring. Shirantha Welikala *et al.* designed an algorithm with 88% monitoring accuracy for non-intrusive load monitoring [14]. The factorial hidden Markov model has been combined with discriminative decomposition sparse for increasing monitoring accuracy [15]. The 90.85% accuracy of load monitoring is obtained by similar time window algorithms [16]. More sample data can be introduced to improve the monitoring accuracy of loads. For example, [17] proposed active-reactive power signals as learning data. Derivative analysis and filter analysis have been proposed to improve classification performance [18]. Reference [19] combined multiple devices from aggregated measurement data for monitoring. A sequencing method has been proposed to reduce the training process of neural networks [20].

An expendable comprehensive load decomposition model has been applied for the load decomposition problem of mixed loads [21]. A spatiotemporal pattern network has been employed for energy disaggregation [22]. A hierarchical hidden Markov model framework with multiple functions has been proposed for non-intrusive load monitoring [22]. The multi-label classification problem has been solved by the currently popular sparse representation-based classification method [23]. A solution called real-time non-intrusive load monitoring has been designed with a high level of recognition accuracy even during the power supply voltage fluctuates sharply [24]. A graph-based multi-label method has an accuracy of 72.85% for load monitoring [25]. The curse of dimensionality can be avoided by multiple sequences of regression learning [26].

The mainly neural networks applied in the references above are roughly based on the following neural networks. The Bayesian method [14] is a classification algorithm based on statistic theories. The Bayesian method has a quite concise structure. The Bayesian method has a speed advantage in classification. However, the Bayesian algorithm does not have a strong classification ability without conditional independence between attributes of things that need to be classified. The hidden Markov model [15] is a graph clustering algorithm. The hidden Markov model described a process of generating an unobservable state sequence from an implicit Markov chain and then generating an observation sequence from the unobservable state sequence. The hidden Markov model is memoryless and cannot apply contextual information. K-nearest neighbors [16] apply Euclidean distance to find the relationship of things that need to be classified. The calculation process of k-nearest neighbors can adapt to the classification of large-scale data. However, the classification accuracy of rare categories is lower than that of rich categories when the number of samples of each category is unbalanced. Deep backpropagation networks [27] are multi-layer feedforward neural networks applied back propagation algorithm. The deep backpropagation networks have a strong ability of nonlinear fitting. Nevertheless, the classification accuracy of the basic deep backpropagation networks cannot meet the need for non-intrusive load monitoring. Convolutional networks [19] extracts features of the image through the action of the convolutional layer and pooling layer. The convolutional networks can accurately classify image features. However, the convolution layer and the pooling layer require complex calculations. The calculation of convolution networks is a hard burden for an embedded device. Long short-term memory networks [26] have the ability to model time series data. The long short-term memory networks do not have the ability to process parallelized data. The training process of the long short-term memory networks is slower than the deep backpropagation networks.

The deficiencies of the above researches are listed as follows.

(1) Most of the above studies employ existing data sets and lack recognition tests on data outside the existing data sets.

(2) Intelligent algorithms have been continuously improved with the continuous development of bionics. The recognition accuracy of the above methods still has promotion space.

(3) The optimization of the data preprocessing level needs to be improved. Real-time multiplication calculation is required for the training process of characteristic data with specific computing power. The multiplication processes of load characteristics are not conducive to the algorithm for load monitoring.

Fast Fourier transforms [9], the operating sounds of the electrical equipment, the step change of the current signal [28], event drivers [29], feature selection [30], and eventual current measurement [31] have been applied for non-intrusive load monitoring. The operating data of loads are applied for non-intrusive load monitoring in this paper. The features of loads can describe loads from multiple dimensions and provide more information for monitoring.

With the continuous development of artificial intelligence technologies in recent years, a flexible transmitter networks method has been proposed as an alternative neuron structure [32]. The flexible transmitter networks entirely refer to the principles of bionics and more deeply simulate the biological characteristics of neurons. Since flexible transmitter networks have a more complex structure with more computing ability, the accuracy of flexible transmitter networks is higher than that of support vector machines and convolutional neural networks [33]. However, the flexible transmitter networks need complex calculations due to the added parameter in the back propagation process of the training process. To maintain a high classification accuracy and reduce the time of the training process of networks, a simple network can be chosen to be combined. In this paper, sampled data is processed through fast Fourier decomposition to be training data. The deep backpropagation networks can satisfy the kind of training data. Apart from satisfaction, the deep backpropagation networks have a quicker back propagation speed than the flexible transmitter networks. The flexible

transmitter networks and the deep propagation networks are combined.

With the continuous reduction of lower performance computing costs and the continuous development of embedded devices, a new generation of information technologies represented by the internet of things [34], cloud computing [35], and edge computing [36] have been widely applied [37]. The preprocessing and monitoring process of loads waveform requires high real-time performance and relatively parallel ability, which is a heavy burden for the computation center. The approach of submerging cloud computing capabilities to the edge of the network close to the client to perform specific computing tasks can greatly enhance computing power and reduce costs [38].

Inspired by deep backpropagation networks, flexible transmitter networks, load features, and edge computing, this paper proposes a deep flexible transmitter networks method to balance monitoring accuracy and monitoring speed for non-intrusive load monitoring. The significant contributions of the article are listed as follows:

(1) The standard loads of home consumers are analyzed and modeled; the generated voltage and current waveforms during the operation of loads are measured; a data set for the operation of the load is formed. Compared with the existing data set, the data set applied in this paper is closer to the real-life situation of loads; the results obtained are more similar to the real-life load application scenarios.

(2) A method is established to preprocess the acquired sampled data. After various sampled data is processed through fast Fourier decomposition, significant parameters are selected for training and recognition. The proposed method improves the monitoring accuracy and does not exceed the acceptable range of computing power demand. After the monitoring process, the proposed method can replace the power decomposition method.

(3) Combining the flexible transmitter networks and the deep backpropagation networks, the proposed algorithm can effectively improve recognition accuracy and can meet the computing requirements of existing computing equipment. After the monitoring process, the proposed algorithm meets the prerequisite of working on embedded devices and has the feasibility of applying to edge computing.

The rest of the paper is presented as follows. The established loads model and the flexible transmitter networks are given in Section II and Section III respectively. Section IV shows case studies. Conclusions are listed in Section V.

II. ESTABLISHMENT OF FEATURE DATA SET FOR LOADS

To obtain the feature data of loads and establish a feature data set of loads. The establishment of the operating data set of the loads model is divided into two parts, i.e., the power supply side and the equipment side.

A. POWER SUPPLY SIDE

A power supply side of real-life projects consists of two parts, i.e., the main excitation part and the noise part (FIGURE 1(a)). The main excitation part is applied to excite the load and to generate the corresponding load waveform. The noise part (FIGURE 1(b)) mainly comes from the influence of interference and power grid clutter when sampling waveform.

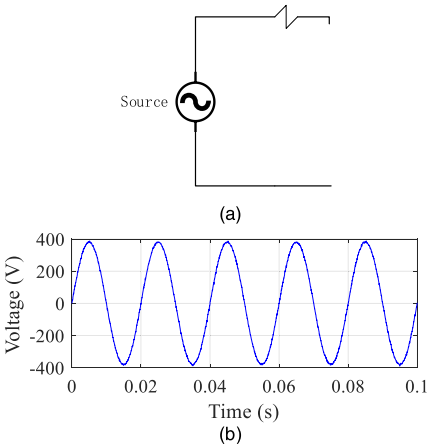

FIGURE 1. Model and voltage curve of generation part: (a) model of generation part; (b) voltage wave of generation part.

B. EQUIPMENT SIDE

In the daily life of power users, roughly three types of equipment often appear, i.e., pure resistance equipment, rotating electrical equipment, and equipment containing power electronics.

1) PURE RESISTANCE EQUIPMENT

The structure of the pure resistance devices is relatively simple. For example, a water heater model comprises a power supply, transformer, and load (FIGURE 2(a)). After setting the resistance, inductance, reactance, and transformer parameters, the waveform of a water heater can be measured directly (FIGURE 2(b)).

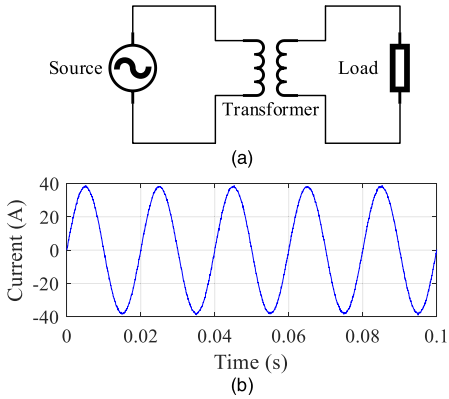

FIGURE 2. Model and current curve of shower heater: (a) model of shower heater; (b) current wave of shower heater.

2) ROTATING ELECTRICAL EQUIPMENT

The fixed-frequency air conditioner, constant frequency fan, and constant frequency refrigerator models in a home consumer are built.

For the fixed-frequency air-conditioning model, the main source of the load is the compressor. The working principle of the air-conditioning compressor is that the motor drives the compressor to perform reciprocating compression in the cavity of a compressor. Therefore, the load torque curve of the motor is set as a triangular wave. The schematic diagram and the measurement waveform of an air conditioning model are shown in FIGURE 3. The letter M in FIGURE 3(a) presents the electrical motor.

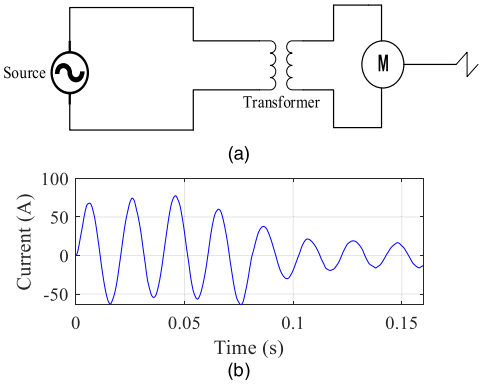

FIGURE 3. Model and current curve of air conditioning: (a) model of air conditioning; (b) current wave of air conditioning.

The fixed-frequency refrigerator is a load that contains a compressor. The model framework is similar to the fixed-frequency air conditioner model. The parameters of the internal transformer and the motor are adjusted as the parameters of the refrigerator. The model and waveform of the fixed frequency refrigerator are shown in FIGURE 4.

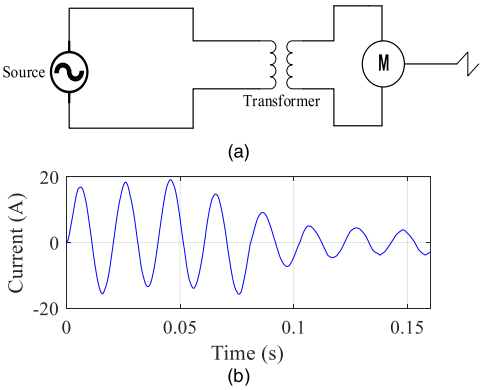

FIGURE 4. Model and current curve of constant frequency refrigerator: (a) model of constant frequency refrigerator; (b) current wave of constant frequency refrigerator.

Fixed frequency fans generally apply alternating current motors and control the fan speed by controlling the voltage (FIGURE 5(a)). The waveform of a fixed frequency fan when the fan runs at full speed is given in FIGURE 5(b).

3) EQUIPMENT CONTAINING POWER ELECTRONICS

Complex power electronic equipment generates different kinds of non-sine waves; the various harmonics contained therein have a certain degree of impact on the power

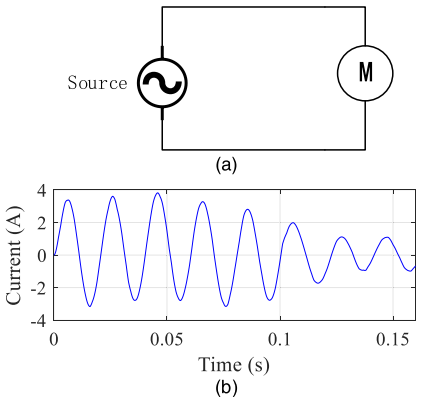

FIGURE 5. Model and current wave of constant frequency fan: (a) model of constant frequency fan; (b) current wave of constant frequency fan.

distribution network. Therefore, non-intrusive load monitoring has tremendous significance for the regular operation of the power grid for monitoring the load during power electronic devices connected to the power grid. Rectifiers and frequency converters are selected to establish the load model of power electronic devices.

For loads with rectifiers, the light-emitting diode lighting system model is established (FIGURE 6(a)). Then, the waveform of the light-emitting diode lighting system is measured as FIGURE 6(b).

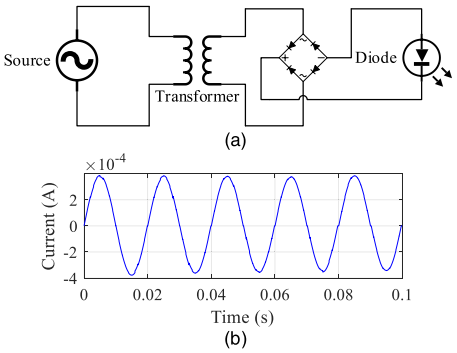

FIGURE 6. Model and current curve of light-emitting diode: (a) model of light-emitting diode; (b) current wave of light-emitting diode.

In the same approach, the washing machine model is established as FIGURE 7(a); the waveform of a washing machine is measured as FIGURE 7(b).

The hairdryer model is established as FIGURE 8(a); the hairdryer waveform is measured as FIGURE 8(b).

For loads containing inverters, inverter air conditioners, inverter fans, and induction cookers are selected as examples. The applied inverters are constructed with the method of rectification-inverter. The rectification part applies a full bridge rectifier circuit composed of thyristors. The inverter part employs an insulated gate bipolar transistor full control bridge inverter circuit (FIGURE 9(a)). The voltage waveform of the inverter part is measured as FIGURE 9(b). The abbreviation of AC in FIGURE 9(a) presents alternative voltage; the letter Q in FIGURE 9(a) presents the power electronic switch.

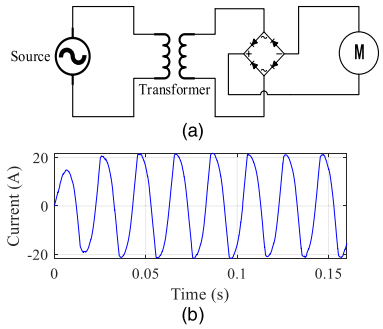

FIGURE 7. Model and current curve of washing machine: (a) model of washing machine; (b) current wave of washing machine.

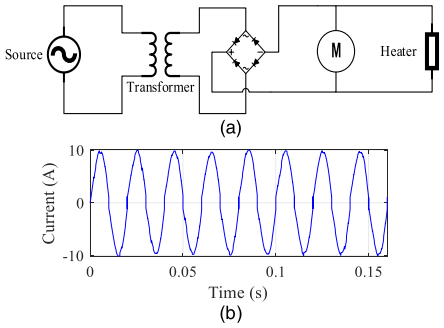

FIGURE 8. Model and current curve of hair dryer: (a) model of hair dryer; (b) current wave of hair dryer.

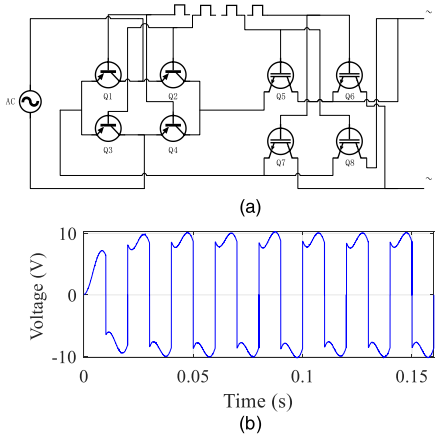

FIGURE 9. Model and voltage wave of frequency converter: (a) model of frequency converter; (b) voltage wave of frequency converter.

The inverter air conditioner model is similar to the fixed frequency air conditioner model when adding an inverter part to control the compressor speed (FIGURE 10(a)). The waveform of the inverter air conditioner is measured as FIGURE 10(b).

When the variable frequency fan model runs at a constant speed, the received load torque is constant. Thus, the load torque of the fan motor is set to a constant value (FIGURE 11(a)). The waveform of the variable frequency fan is measured as FIGURE 11(b). The abbreviation of Cons in FIGURE 11(b) presents constant.

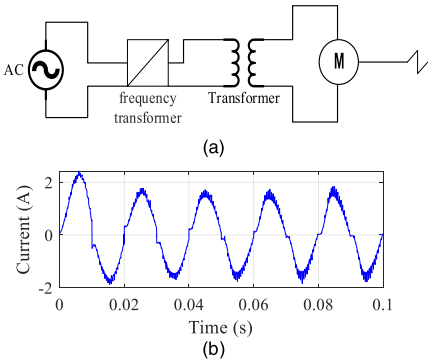

FIGURE 10. Model and current curve of inverter air conditioner: (a) model of inverter air conditioner; (b) current wave of inverter air conditioner.

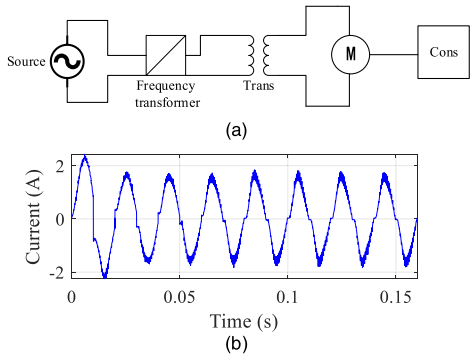

FIGURE 11. Model and current curve of variable frequency fan: (a) model of variable frequency fan; (b) current wave of variable frequency fan.

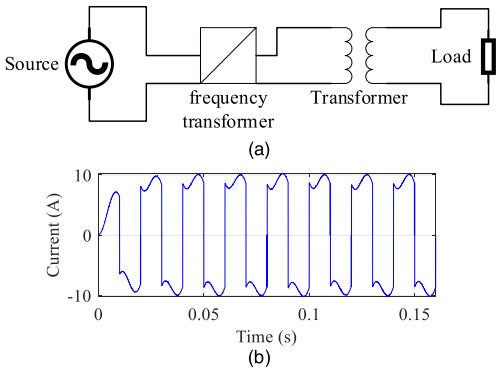

FIGURE 12. Model and current curve of induction cooker: (a) model of induction cooker; (b) current wave of induction cooker.

The working principle of the induction cooker model is that: generate a high-frequency alternating current by a frequency converter; then, a high-frequency magnetic field is generated through a coil on the bottom of the pot for generating heat. Therefore, the heating process of the induction cooker can be equivalent to a model in which a frequency converter drives a transformer. The transformer of the induction cooker is connected by an equivalent resistance load (FIGURE 12(a)). The waveform of the induction cooker is measured as FIGURE 12(b).

C. SAMPLE DATA OF ESTABLISHING FEATURE DATA SET OF LOADS

To ensure the accuracy of the acquired waveform and the perception of small fluctuations in the process of acquiring

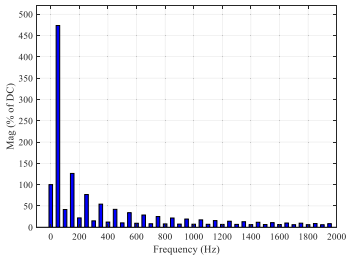

FIGURE 13. Frequency spectrum of electromagnetic furnace.

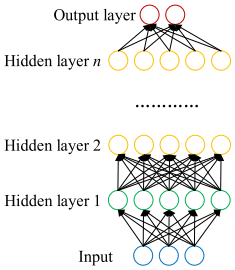

FIGURE 14. Schematic diagram of deep neural networks.

waveform data, the sampling frequency of the sampled oscilloscope is set to be 10 kHz. Therefore, the waveform data of each period consists of thousands of discrete time-series data points. If the raw data is directly inputted into the neural networks for training, the internal structure of the neural networks should be designed large enough for accepting such high-order input data. To achieve the faster speed and higher fitting effect level of the neural networks, preprocessing the input data with reducing the order to an acceptable level is necessary to establish the neural networks.

Considering that loads run on the sine wave excitation of power systems, the generated current after the excitation has prominent sinusoidal characteristics. Based on the characteristics, the harmonics of the load current can be extracted as the feature of loads, set as the monitoring feature of the deep flexible transmitter networks.

A discrete Fourier transform is applied for converting the discrete sampled current and voltage data to extract the harmonic characteristics of the current waveform of loads. The induction cooker model, which has the most severe current waveform distortion, is selected to describe the converted process. After one period of the waveform of the induction cooker model is converted by the Fourier transform, a converted frequency spectrum of the induction cooker is formed as FIGURE 13.

After the waveform of loads is performed by Fourier transform, the discrete time-domain waveform of the induction cooker is transformed into a frequency spectrum (FIGURE 13). The frequency spectrum distribution shows that the fundamental wave still accounts for most of the induction cooker waveform. Furthermore, the frequency spectrum of the induction cooker shows that the distribution value of the frequency spectrum of the frequency that higher than 400 Hz is lesser than half of the maximum distribution value of the frequency spectrum of the induction cooker. Therefore, the distribution values of the frequency spectrum of system frequency that lower than 400 Hz are selected as the training data of the proposed method. The amplitude and phase angle that performed the Fourier transform of the full simulation period of loads are stored as the training data of loads feature.

III. DEEP BACKPROPAGATION NETWORKS

A shallow neural networks method consists of an input, hidden, and output layer. The hidden layer sums the weighted sum of the input layer by a weight matrix. Then, the result of the output layer can be obtained by a nonlinear activation function. The mathematical expression of shallow neural networks is listed as follows:

$$\boldsymbol{a}^{l+1} = f(\boldsymbol{W}^l \boldsymbol{a}^l + \boldsymbol{b}^l) \tag{1}$$

where $\boldsymbol{a}^{l+1}$ and $\boldsymbol{a}^l$ are the output matrices of the $(l+1)$-th and l-th hidden layers, respectively; $f(\cdot)$ is the activation function of shallow neural networks; $\boldsymbol{W}^l$ is the weight matrix of the l-th layer; $\boldsymbol{b}^l$ is the deviation value matrix of the l-th layer of shallow neural networks

Deep neural networks (FIGURE 14), which expand the number of hidden layers from shallow networks, apply the output of the low-level hidden layer as the input of the high-level hidden layer. Multiple hidden layers enhance the characterization degree of the nonlinearity of the deep neural networks. Therefore, deep neural networks can fit nonlinear data more accurately and have higher generalization performance and robustness than shallow neural networks.

A. FLEXIBLE TRANSMITTER NETWORKS

The flexible transmitter networks have been improved by Dr. Shaoqun Zhang *et al.* from the model proposed by McCulloch and Pitts [32]. The weight nonlinearity matrix processes the information transmitted by the neurons in the upper layer of the McCulloch-Pitts model. Then, the information transmitted by the neurons in the next layer is generated under the action of an activation function. Compared with the McCulloch-Pitts model, the flexible transmitter networks introduce a novel bionic principle. When constructing neurons, the role of neuron axons and dendrites are considered at the same time. Besides, the number of the feedforward channel parameters of the flexible transmitter networks is increased to two (FIGURE 15). Apart from the number of parameters of the feedforward process, the flexible transmitter networks introduced the third parameter to simulate the memory strength of neurons. Since the flexible transmitter networks are more similar to biological neuron characteristics, the flexible transmitter networks have a higher fitting accuracy than the McCulloch-Pitts model.

The neuron framework of the flexible transmitter networks includes two feedforward channel parameters (W, V) and an iterative memory strength parameter M. The training process of the flexible transmitter networks includes the following steps in detail.

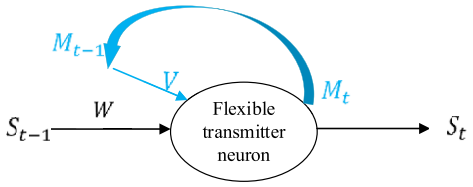

FIGURE 15. Schematic diagram of neurons in flexible transmitter networks.

(1) The feedforward process of the flexible transmitter networks. With the t-th neuron of the flexible transmitter networks are selected as an example, the feedforward channel of the t-th neuron of the flexible transmitter networks has two parameters (W, V). The neuron memory strength of the t-th neuron of the flexible transmitter networks has an iterative parameter M. The feedforward function of the flexible transmitter networks is described in complex numbers, as follows [32],

$$S_t^l + M_t^l i = \sigma \left[\begin{array}{l} \left(\alpha W^l S_t^{l-1} - \beta V^l M_{t-1}^l\right) \\ + \left(\beta W^l S_t^{l-1} + \alpha V^l M_{t-1}^l\right) i \end{array} \right] \quad (2)$$

where W^l and V^l are the feedforward weight matrix and memory weight matrix of the l-th layer of the flexible transmitter networks, respectively; S_t^l is the output of the current iteration of the flexible transmitter networks; S_{t-1}^l is the input of the current iteration of the flexible transmitter networks; M_{t-1}^l is the memory strength that calculated in the current iteration; M_t^l is applied in the next iteration; σ is the excitation function such as the sigmoid function; the excitation function has two parts, i.e., a real part function and an imaginary part function; both α and β are adjustable parameters in the real number domain of the flexible transmitter networks.

(2) The backpropagation process of the flexible transmitter networks. After completing each iteration of the flexible transmitter networks, the parameters of the flexible transmitter networks are corrected according to the iteration results. Before calculating the gradient of the flexible transmitter networks, define [32]:

$$Real = \alpha W^l S_t^{l-1} - \beta V^l M_{t-1}^l, \quad (3)$$

$$Imag = \beta W^l S_t^{l-1} + \alpha V^l M_{t-1}^l, \quad (4)$$

where *Real* and *Imag* are the real and imaginary parts of the parameter of the feedforward process of the flexible transmitter networks, respectively.

The function of the backpropagation process of the flexible transmitter networks can be derived as [32]:

$$\begin{aligned} &\left(\nabla_{W^l} E, \nabla_{V^l} E\right) \\ &= \int_{t=1}^{T} \left(\nabla W_t^l, \nabla V_t^l\right) dt \\ &= \int_{t=1}^{T} \left(\delta S_t^l \odot \frac{\partial S_t^l}{\partial Real_t^l}\right) \cdot \left(\frac{\partial Real_t^l}{\partial W^l}, \frac{\partial Real_t^l}{\partial V^l}\right) dt \end{aligned} \quad (5)$$

where $\nabla_W E$ and $\nabla_V E$ are the backpropagation gradients of the weight matrices of the flexible transmitter networks W and V, respectively; ∇W_t^l and ∇V_t^l represent the gradient of the backpropagation of the l-th layer of the flexible transmitter networks in the t-th iteration, respectively; δS_t^l is the backpropagation correction difference calculated in the training process of the flexible transmitter networks; $\odot$ is the dot multiplication operation; $\frac{\partial S_t^l}{\partial Real_t^l}$ is the derivative vector of the activation function $\left[\sigma'_{real}\left(Real_t^l|_1\right), \ldots, \sigma'_{real}\left(Real_t^l|_{n_l}\right)\right]^T$ applied in the training process of the flexible transmitter networks; $\frac{\partial Real_t^l}{\partial W^l}$ and $\frac{\partial Real_t^l}{\partial V^l}$ are the complex backpropagation calculation method of backpropagation process of the flexible transmitter networks for these two backpropagation pipelines related to W_t^l and V_t^l, respectively; where [32]:

$$\delta S_t^l = \alpha \cdot \left(W^L\right)^T \cdot \left(\delta S_t^{l+1} \odot \sigma'\left(Real_t^{l+1}\right)\right) \quad (6)$$

$$\begin{aligned} \frac{\partial Real_t^l}{\partial W_{ij}^l} &= \begin{pmatrix} \Delta W_{11}^l & \cdots & \Delta W_{1n_{l-1}}^l \\ \vdots & \ddots & \vdots \\ \Delta W_{n_l 1}^l & \cdots & \Delta W_{n_l n_{l-1}}^l \end{pmatrix} \\ \frac{\partial Real_t^l}{\partial V_{ik}^l} &= \begin{pmatrix} \Delta V_{11}^l & \cdots & \Delta V_{1n_{l-1}}^l \\ \vdots & \ddots & \vdots \\ \Delta V_{n_l 1}^l & \cdots & \Delta V_{n_l n_{l-1}}^l \end{pmatrix} \end{aligned} \quad (7)$$

where [32]:

$$\begin{aligned} \Delta W_{ij}^l &= \alpha \cdot S_t^l|_j - \beta \cdot \sum_{k=1}^{n_l} V_{ik}^l \frac{\partial M_{t-1}^l|_k}{\partial W_{ij}^l} \\ \Delta V_{ik}^l &= -\beta \left[M_{t-1}^l|_k + \sum_{k=1}^{n_l} V_{ik}^l \frac{\partial M_{t-1}^l|_k}{\partial V_{ij}^l} \right] \end{aligned} \quad (8)$$

where [32]:

$$\begin{cases} \dfrac{\partial M_t^l|_k}{\partial W_{ij}^l} = \sigma' \cdot \dfrac{\partial Imag_t^l|_k}{\partial W_{ij}^l} \\ = \begin{cases} \sigma'\left(Imag_t^l|_k\right) \cdot \left[\begin{array}{l} \beta \cdot S_t^{l-1}|_j \\ + \alpha \cdot \sum_{h=1}^{n_l} \tilde{V}_{ih} \dfrac{\partial M_{t-1}^l|_h}{\partial W_{ij}^l} \end{array} \right], & i = k; \\ 0, & i \neq k. \end{cases} \\ \dfrac{\partial M_t^l|_k}{\partial V_{ik}^l} = \sigma' \cdot \dfrac{\partial Imag_t^l|_k}{\partial V_{ik}^l} \\ = \begin{cases} \sigma'\left(Imag_t^l|_k\right) \cdot \alpha \cdot \left[\begin{array}{l} M_{t-1}^l|_k \\ + \sum_{j=1}^{n_l} \tilde{V}_{kj}^l \dfrac{\partial M_{t-1}^l|_j}{\partial V_{kk}^l} \end{array} \right], & i = k; \\ \sigma'\left(Imag_t^l|_k\right) \cdot \alpha \cdot \sum_{j=1}^{n_l} \tilde{V}_{kj}^l \dfrac{\partial M_{t-1}^l|_j}{\partial V_{ik}^l}, & i \neq k. \end{cases} \end{cases} \quad (9)$$

With the calculation of (5), both W and V are corrected with a given fixed learning rate η.

$$\begin{cases} \hat{W} = W + \eta \cdot \nabla_{W^l} E \\ \hat{V} = V + \eta \cdot \nabla_{V^l} E \end{cases} \tag{10}$$

The loss function E is calculated after the backpropagation process if the flexible transmitter networks are over. The loss function E is the integral loss function $E(W, V)$ of the flexible transmitter networks, which fits the loss by the square of the difference between the real value of the given training set and the predicted value of the flexible transmitter networks, as follows [32]:

$$E(W, V) = \frac{1}{2} \int_{t=1}^{T} \sum_{i=1}^{n_l} \left(Y_t^{\text{true}}|_i - Y_t|_i \right)^2 dt \tag{11}$$

where Y_t is the output signal of the final layer of flexible transmitter networks.

According to the loss function calculation results (11), whether the training process is finished can be judged. The last iteration M is calculated after reaching the expected index [32].

$$M_t = \sigma \left(\beta \hat{W} X_t + \alpha \hat{V} M_{t-1} \right). \tag{12}$$

After the training process of the flexible transmitter networks is over, the neurons of the flexible transmitter networks can solve regression or classification problems.

B. DEEP FLEXIBLE TRANSMITTER NETWORKS

Two layers of cumulative sum calculation are required when calculating the backpropagation error of the flexible transmitter networks. Therefore, the time complexity of calculating the backpropagation error of the flexible transmitter networks is $O(2mn^2)$, where m and n are the numbers of neurons in the layer which calculating the backpropagation error of the flexible transmitter networks and the previous layer, respectively. When the neurons of the flexible transmitter networks with multi-layer are superimposed to work or when the number of neurons in a single layer of the flexible transmitter networks is increased, the calculation complexity of the flexible training process transmitter networks is increased. The increased calculation effectively influences the speed of the flexible transmitter networks. However, if the number of the layers of the flexible transmitter networks or the number of neurons in a single layer of the flexible transmitter networks is reduced due to the pursuit of the training process speed of the flexible transmitter networks, the fitting accuracy of the flexible transmitter networks is reduced; consequently, the regression accuracy or recognition accuracy of the flexible transmitter networks is decreased. Although the fitting accuracy of the deep backpropagation networks is lower than that of flexible transmitter networks, the time complexity of deep backpropagation networks with backpropagation calculation is $O(n)$. Compared with the training time of flexible transmitter networks, the deep backpropagation networks need a shorter training time. With the advantage of the accuracy of the flexible transmitter networks and the advantage of the speed of deep backpropagation networks, a deep flexible transmitter network is established in this paper (FIGURE 16).

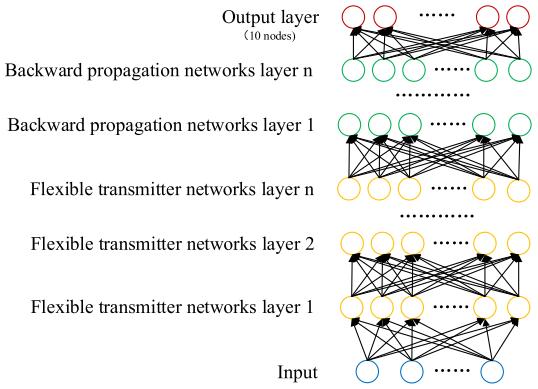

FIGURE 16. Schematic diagram of deep flexible transmitter networks.

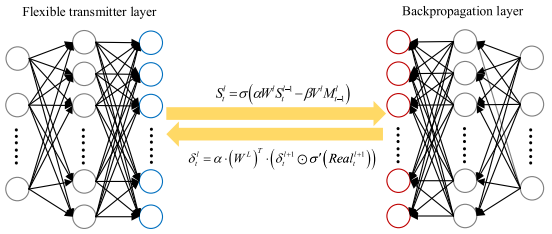

FIGURE 17. Backpropagation process of deep flexible transmitter networks.

$$\delta^l = (W^{l+1})^T \delta^{l+1} \odot \sigma'(z^l) \tag{13}$$

where δ^l is the backpropagation error of the l-th layer of the deep backpropagation networks; W is the weight matrix of the deep backpropagation networks; z^l is the output of the l-th layer of the deep backpropagation networks.

The gradient descent method of the McCulloch-Pitts model, which is the same as the method applied in the deep backpropagation networks, is applied to calculate the backpropagation parameters of the training process of the flexible transmitter networks. The error transfer function of the backpropagation process of the training process of the transmitter networks is listed as follows [32]:

The error transfer function of the backpropagation process of the deep backpropagation networks is described as:

$$\delta_t^l = \alpha \cdot \left(W^L \right)^T \cdot \left(\delta_t^{l+1} \odot \sigma' \left(Real_t^{l+1} \right) \right). \tag{14}$$

The backpropagation method of a single backpropagation variable is applied in both the deep backpropagation and flexible transmitter networks. During the backpropagation process of the deep flexible transmitter networks, the backpropagation error of the first layer of the deep backpropagation networks is transmitted to the last layer of the flexible transmitter networks. Two-layer flexible transmitter networks and two-layer backpropagation networks are combined as FIGURE 17.

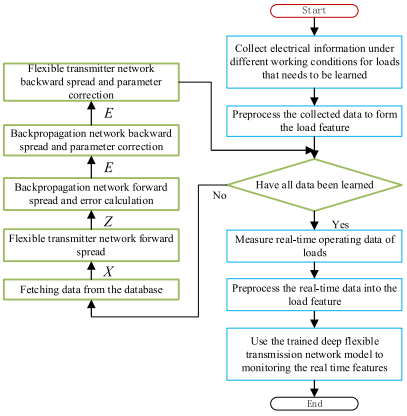

FIGURE 18. Flowchart of the algorithm of deep flexible transmitter networks for non-intrusive load monitoring.

C. NON-INTRUSIVE LOAD MONITORING APPLYING DEEP FLEXIBLE TRANSMITTER NETWORKS

After obtaining the load feature, the loads feature data is divided into the training set and test set. The flow chart of the deep flexible transmitter networks for non-intrusive load monitoring is shown in FIGURE 18, where X is the input data; Z is the output data; E is the error between output data and the real data.

Each of the five periods of the loads feature data is set as a group; one data of each group is selected as the test data. All the test data is stored in the test set; the rest data is configured as the training set. Deep flexible transmitter networks are established to train the training set mentioned above. After the training process is completed, the test set mentioned above is applied to test the recognition accuracy of the trained deep flexible transmitter networks.

IV. CASES STUDIES

The hardware configuration applied in the learning process in this experiment is a Mac: the system is macOS 11.0; the processor is Intel (R) Core (TM) i7-6820HQ; the main frequency is 2.70 GHz; the memory is 16 GB. The hardware configuration applied in the monitoring process of this experiment is ARM: the system is Linux; the processor is Cortex-A53 architecture; the main frequency is 1.2 GHz. A total of 128 M of memory is allocated for the program to simulate the scenario of running on embedded devices or edge computing devices (FIGURE 19). The hardware system consists of a voltage sensor, a current sensor, and an ARM processor. The hardware system can obtain voltage and current data from the sensor. Then, the monitoring results can be displayed on the screen and can be uploaded to the cloud.

A. EXPERIMENTS

The waveform of these loads mentioned in Section II is measured. The sampling frequency of 10 kHz is applied to sample these loads with a total of 5 s. After Fourier transform is applied to convert the sampled data of the current and voltage

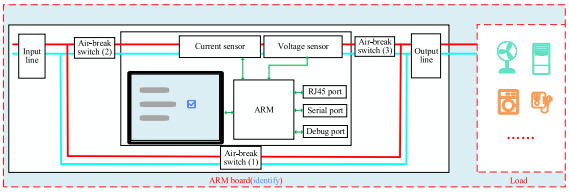

FIGURE 19. Schematic diagram of recognition equipment.

information of loads, the converted data are stored in the data set.

In this experiment, the established deep flexible transmitter networks are composed of two layers of the flexible transmitter networks and five layers of the backpropagation neural networks (FIGURE 20).

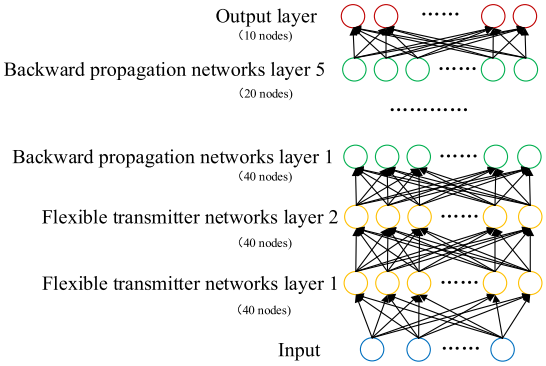

FIGURE 20. Structure diagram of deep flexible transmitter networks.

The iterations-accuracy of the training process of the deep flexible transmitter networks is illustrated as FIGURE 21. The training set applied in the training process of the deep flexible transmitter networks consists of 2000 data. The mini-batch parameter is set to 32; the epochs number is set to 60 during the training process of the flexible transmitter networks; thus, the training process of the flexible transmitter networks consists of 3720 iterations. In the training process of

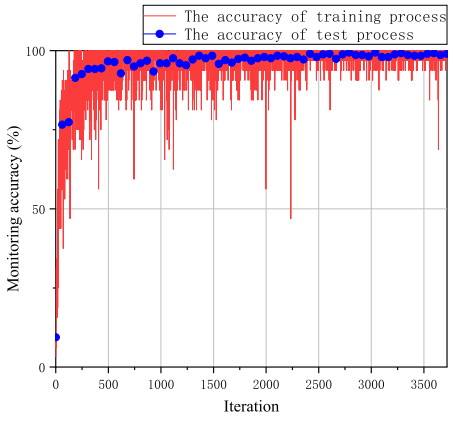

FIGURE 21. Recognition accuracy of training process and testing process of deep flexible transmitter networks.

the deep flexible transmitter networks, the test set is applied to verify the monitoring accuracy after each epoch (blue point in FIGURE 21).

After 500 iterations, the maximum recognition accuracy of the training process of the deep flexible transmitter networks is 100%. The monitoring accuracy of the deep flexible transmitter networks reached 99.00% in the final verification after the training process is completed. The training process of the deep flexible transmitter networks shows that the deep flexible transmitter networks are effective and have a strong learning ability.

Five networks (i.e., deep backpropagation networks, support vector machines, convolutional neural networks, recurrent neural networks, and long short-term memory networks) are trained with the same training set as the deep flexible transmitter networks. The test set employed in the deep flexible transmitter networks is applied to test the monitoring accuracy of these five networks. In the training process of these five networks, the same learning rate parameters and mini-batch learning parameters of the training process of the deep flexible transmitter networks are set to these five networks. The following conclusions can be illustrated by comparing the training process of these five networks and deep flexible transmitter networks:

(1) For monitoring accuracy, the comparison between deep flexible transmitter networks and these five networks in the training process is shown in FIGURE 22.

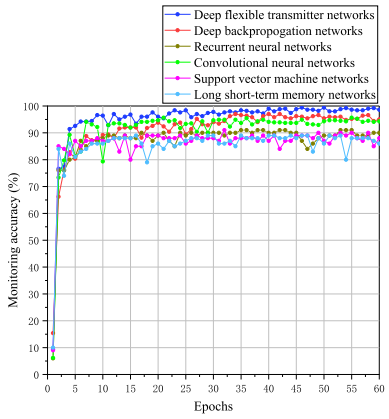

FIGURE 22. Comparison of deep flexible transmitter networks and other networks.

The monitoring accuracy of the training process of deep flexible transmitter networks and these five networks (FIGURE 22) shows that the recognition accuracy of the deep flexible transmitter networks is higher than that of these five networks in the final epoch. For the speed of the training process, the deep flexible transmitter networks have a higher rate than these five networks. Specifically, in the fifth epoch, the deep flexible transmitter networks have reached a monitoring accuracy of 90.00%. In contrast, the deep backpropagation networks can only achieve this training effect at the 15th epoch, and the convolutional neural networks can only maintain this training effect at the seventh epoch. Therefore, in terms of monitoring accuracy, the performance of the deep flexible transmitter networks is relatively excellent.

(2) In terms of training time, affected by the complex derivation process of the backpropagation process of the training process of the flexible transmitter networks, the training process of the deep flexible transmitter networks lasted 42 minutes. The deep backpropagation networks and the convolutional neural networks, which have an accuracy of more than 95.00% in the training process, are selected as comparison networks. The training process of the deep backpropagation networks lasted 18 minutes and 9 s. The training process of the convolutional networks lasted 67 minutes and 23 s. The result of the training process of the deep flexible transmitter networks, the deep backpropagation networks, and the convolutional networks shows that the speed of the training process of the deep flexible transmitter networks is about half of that of the deep backpropagation networks; the speed of the training process of convolutional neural networks is lower than that of deep flexible transmitter networks.

After the training process of the deep flexible transmitter networks, the parameters of the trained deep flexible transmitter networks can be imported into the monitoring equipment for monitoring the waveform of the loads. A total of five waveform segments in the test set are randomly selected for monitoring. The time-consuming monitoring process of the deep flexible transmitter networks and deep backpropagation networks for these five waveform segments is 0.0042, 0.0067, 0.0077, 0.0034, and 0.0059 s.

The time-consuming process of the monitoring process is less than one electrical period (i.e., 0.02 s) when the proposed monitoring method is running on the embedded device. Therefore, the monitoring process of the deep flexible transmitter networks is within the allowable range of the computing power of embedded devices and edge computing devices.

The monitoring accuracies of the flexible transmitter networks and these five networks are given in Table 1.

TABLE 1. Monitoring accuracy of deep flexible transmitter networks and other networks.

Type of neural networks	Recognition accuracy	Training time
Deep flexible transmitter networks	99.00%	42 minutes 1 s
Deep backpropagation networks	94.00%	18 minutes 9 s
Support vector machines	88.00%	18 minutes 32 s
Convolutional neural networks	93.51%	67 minutes 23 s
Recurrent neural networks	90.00%	27 minutes 87 s
Long short-term memory networks	86.00%	19 minutes 42 s

Compared with the five networks, the features of the deep flexible transmitter networks can be summarized as follows:

(1) The monitoring accuracy of the deep flexible transmitter networks dramatically exceeds that of the five

networks. The deep backpropagation networks, which have the highest monitoring accuracy in the five networks, are select as the main comparison networks. The original data of the test results of the deep flexible transmitter networks and the deep backpropagation networks show that the lower monitoring accuracy obtained by the deep backpropagation networks is mainly from the initial startup time of the motor-containing loads. The measured waveform of the motor-containing loads shows that the current waveform is not stable during the startup period. The measured waveform contains many harmonic waves and aperiodic waves. The monitoring results show that: the deep backpropagation networks achieve lower learning and monitoring effects on irregular periods, while the deep flexible transmitter networks can monitor the initial moment of the motor startup more accurately. With improving monitoring ability at initial startup time, the proposed algorithm can recognize the load in as little calculation time as possible. The speed of the monitoring system to monitor the load can be significantly increased.

(2) The flexible transmitter networks are slower than the deep backpropagation networks, the support vector machines, the recurrent neural networks, and long-term memory networks in the training process. However, the support vector machines, the recurrent neural networks, and the long short-term memory networks cannot exceed the level of 90% in monitoring accuracy. Although the deep flexible transmitter networks are slower than the deep backpropagation networks in the training process, the heavy computing requirement appears in the backpropagation process of the training process of the deep flexible transmitter networks. In the forward propagation process of the test process of the deep flexible transmitter networks, the calculation method of the deep flexible transmitter networks is similar to that of the deep backpropagation networks. Therefore, after the training process of the deep flexible transmitter networks is completed, non-intrusive load monitoring no longer demands heavy computing power. Consequently, the calculation of the forward propagation of the deep flexible transmitter networks does not require much calculation time. Based on the calculation method of the deep flexible transmitter networks in the forward propagation process, the deep flexible transmitter networks require the same computing power as the deep backpropagation networks in the recognition process. After testing the deep flexible transmitter networks, the recognition work based on the deep flexible transmitter networks can be deployed on embedded devices.

B. DISCUSSIONS

Numerous adjustments are needed for the network structure of the deep flexible transmitter networks in this paper. During the adjustment process, the monitoring accuracy of the deep flexible transmitter networks can be effectively improved by increasing the number of layers of the flexible transmitter networks. However, when the number of layers of the flexible transmitter networks is increased to larger than 3, the recognition accuracy of the deep flexible transmitter networks is no longer apparent. Moreover, increasing the number of layers of the flexible transmitter networks reduces the speed of the training process of the deep flexible transmitter networks. Another way to improve monitoring accuracy is to expand the number of neurons in each layer of the flexible transmitter networks. However, when the number of neurons in a single layer of the deep flexible transmitter networks exceeds 50, the improvement of the monitoring accuracy of the deep flexible transmitter networks is no longer apparent. However, the unexpected loss of monitoring accuracy of the deep flexible transmitter networks may even occur. Besides, the effect of expanding the number of neurons of a single layer of the part of the deep flexible transmitter networks is more severe than increasing the layers of the part of the flexible transmitter networks on the speed. Therefore, for real-life problems, both the number of layers and the number of neurons in a single layer of the deep flexible transmitter networks should be selected reasonably.

During the training process of the proposed algorithm, various zero elements appear in the error matrix during the backpropagation process. Besides, the appearance of zero elements has a particular arrangement law. In further research, the generation rules of zero elements could be directly generated in the backpropagation process of the training process of the deep flexible transmitter networks. The generation rules of the backpropagation operation of the deep flexible transmitter networks can save training time and improve the performance of the training process.

V. CONCLUSION

This paper proposes deep flexible transmitter networks to monitor loads of power distribution networks. The feasibility and effectiveness of the proposed algorithm are verified under a hardware system. The significant characteristics of the proposed approach are listed as follows.

(1) To improve the speed of non-intrusive load monitoring, the fast Fourier transform and the reasonable selection of the harmonic order reduce the amount of data are configured as the measured data. The proposed approach has a higher monitoring accuracy under a small architecture and achieves a high training and a high monitoring speed by reducing the amount of data that needs to be monitored.

(2) To improve non-intrusive load monitoring accuracy, this paper introduces flexible transmitter networks into deep learning. To mitigate the negative influence on the training speed of the deep flexible transmitter networks, the flexible transmitter networks, and the deep backpropagation networks are fused. The advantages of these two types of networks are combined to achieve high monitoring accuracy with ensuring the calculation requirement of the deep flexible transmitter networks is maintained within an acceptable range for the training process and the monitoring process.

(3) The experiment of deep backpropagation networks, deep flexible transmitter networks, support vector machines,

convolutional neural networks, recurrent neural networks, and long short-term memory networks shows that the deep flexible transmitter networks can achieve non-intrusive load monitoring more accurately. Besides, the training time of the proposed approach is within a reasonable time range. The proposed approach meets the requirements of embedded devices and edge computing devices.

In future researches, (i) the theory of increment learning could be added to the deep flexible transmitter networks for the addition of more types of loads at any time; (ii) more simple calculation and optimization processes of the backpropagation process could be designed to increase the monitoring accuracy, the speed of the training process, and the monitoring process of the deep flexible transmitter networks.

REFERENCES

[1] Y. Sun, X. Xie, Q. Wang, L. Zhang, Y. Li, and Z. Jin, "A bottom-up approach to evaluate the harmonics and power of home appliances in residential areas," *Appl. Energy*, vol. 259, Feb. 2020, Art. no. 114207, doi: 10.1016/j.apenergy.2019.114207.

[2] X. Cao, J. Wang, and B. Zeng, "A study on the strong duality of second-order conic relaxation of AC optimal power flow in radial networks," *IEEE Trans. Power Syst.*, early access, Jun. 9, 2021, doi: 10.1109/TPWRS.2021.3087639.

[3] D. Garcia-Perez, D. Perez-Lopez, I. Diaz-Blanco, A. Gonzalez-Muniz, M. Dominguez-Gonzalez, and A. A. C. Vega, "Fully-convolutional denoising auto-encoders for NILM in large non-residential buildings," *IEEE Trans. Smart Grid*, vol. 12, no. 3, pp. 2722–2731, May 2021, doi: 10.1109/TSG.2020.3047712.

[4] X. Cao, T. Cao, F. Gao, and X. Guan, "Risk-averse storage planning for improving RES hosting capacity under uncertain siting choice," *IEEE Trans. Sustain. Energy*, early access, Apr. 27, 2021, doi: 10.1109/TSTE.2021.3075615.

[5] K. Wang, H. Li, S. Maharjan, Y. Zhang, and S. Guo, "Green energy scheduling for demand side management in the smart grid," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 2, pp. 596–611, Jun. 2018, doi: 10.1109/TGCN.2018.2797533.

[6] H. Rashid, P. Singh, V. Stankovic, and L. Stankovic, "Can non-intrusive load monitoring be used for identifying an appliance's anomalous behaviour?" *Appl. Energy*, vol. 238, pp. 796–805, Mar. 2019, doi: 10.1016/j.apenergy.2019.01.061.

[7] L. Nolting, T. Spiegel, M. Reich, M. Adam, and A. Praktiknjo, "Can energy system modeling benefit from artificial neural networks? Application of two-stage metamodels to reduce computation of security of supply assessments," *Comput. Ind. Eng.*, vol. 142, Apr. 2020, Art. no. 106334, doi: 10.1016/j.cie.2020.106334.

[8] T. Abtahi, C. Shea, A. Kulkarni, and T. Mohsenin, "Accelerating convolutional neural network with FFT on embedded hardware," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 26, no. 9, pp. 1737–1749, Sep. 2018, doi: 10.1109/TVLSI.2018.2825145.

[9] Q. Liu, K. M. Kamoto, X. Liu, M. Sun, and N. Linge, "Low-complexity non-intrusive load monitoring using unsupervised learning and generalized appliance models," *IEEE Trans. Consum. Electron.*, vol. 65, no. 1, pp. 28–37, Feb. 2019, doi: 10.1109/TCE.2019.2891160.

[10] V. Singhal, J. Maggu, and A. Majumdar, "Simultaneous detection of multiple appliances from smart-meter measurements via multi-label consistent deep dictionary learning and deep transform learning," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2969–2978, May 2019, doi: 10.1109/TSG.2018.2815763.

[11] K. He, D. Jakovetic, B. Zhao, V. Stankovic, L. Stankovic, and S. Cheng, "A generic optimisation-based approach for improving non-intrusive load monitoring," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6472–6480, Nov. 2019, doi: 10.1109/TSG.2019.2906012.

[12] W. Kong, Z. Y. Dong, B. Wang, J. Zhao, and J. Huang, "A practical solution for non-intrusive type II load monitoring based on deep learning and post-processing," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 148–160, Jan. 2020, doi: 10.1109/TSG.2019.2918330.

[13] S. Houidi, F. Auger, H. B. A. Sethom, D. Fourer, and L. Miègeville, "Multivariate event detection methods for non-intrusive load monitoring in smart Homes and residential buildings," *Energy Buildings*, vol. 208, Feb. 2020, Art. no. 109624, doi: 10.1016/j.enbuild.2019.109624.

[14] S. Welikala, C. Dinesh, M. P. B. Ekanayake, R. I. Godaliyadda, and J. Ekanayake, "Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 448–461, Jan. 2019, doi: 10.1109/TSG.2017.2743760.

[15] B. Zhao, M. Ye, L. Stankovic, and V. Stankovic, "Non-intrusive load disaggregation solutions for very low-rate smart meter data," *Appl. Energy*, vol. 268, Jun. 2020, Art. no. 114949, doi: 10.1016/j.apenergy.2020.114949.

[16] X. Shi, H. Ming, S. Shakkottai, L. Xie, and J. Yao, "Nonintrusive load monitoring in residential households with low-resolution data," *Appl. Energy*, vol. 252, Oct. 2019, Art. no. 113283, doi: 10.1016/j.apenergy.2019.05.086.

[17] R. Bonfigli, E. Principi, M. Fagiani, M. Severini, S. Squartini, and F. Piazza, "Non-intrusive load monitoring by using active and reactive power in additive factorial hidden Markov models," *Appl. Energy*, vol. 208, pp. 1590–1607, Dec. 2017, doi: 10.1016/j.apenergy.2017.08.203.

[18] M. Lu and Z. Li, "A hybrid event detection approach for non-intrusive load monitoring," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 528–540, Jan. 2020, doi: 10.1109/TSG.2019.2924862.

[19] F. Ciancetta, G. Bucci, E. Fiorucci, S. Mari, and A. Fioravanti, "A new convolutional neural network-based system for NILM applications," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–12, Nov. 2020, Art. no. 1501112, doi: 10.1109/TIM.2020.3035193.

[20] Z. Jia, L. Yang, Z. Zhang, H. Liu, and F. Kong, "Sequence to point learning based on bidirectional dilated residual network for non-intrusive load monitoring," *Int. J. Electr. Power Energy Syst.*, vol. 129, Jul. 2021, Art. no. 106837, doi: 10.1016/j.ijepes.2021.106837.

[21] T. Y. Ji, L. Liu, T. S. Wang, W. B. Lin, M. S. Li, and Q. H. Wu, "Non-intrusive load monitoring using additive factorial approximate maximum a posteriori based on iterative fuzzy C-means," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6667–6677, Nov. 2019, doi: 10.1109/TSG.2019.2909931.

[22] C. Liu, A. Akintayo, Z. Jiang, G. P. Henze, and S. Sarkar, "Multivariate exploration of non-intrusive load monitoring via spatiotemporal pattern network," *Appl. Energy*, vol. 211, pp. 1106–1122, Feb. 2018, doi: 10.1016/j.apenergy.2017.12.026.

[23] S. Singh and A. Majumdar, "Non-intrusive load monitoring via multi-label sparse representation-based classification," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1799–1801, Mar. 2020, doi: 10.1109/TSG.2019.2938090.

[24] S. Welikala, N. Thelasingha, M. Akram, P. B. Ekanayake, R. I. Godaliyadda, and J. B. Ekanayake, "Implementation of a robust real-time non-intrusive load monitoring solution," *Appl. Energy*, vol. 238, pp. 1519–1529, Mar. 2019, doi: 10.1016/j.apenergy.2019.01.167.

[25] D. Li and S. Dick, "Residential household non-intrusive load monitoring via graph-based multi-label semi-supervised learning," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4615–4627, Jul. 2019, doi: 10.1109/TSG.2018.2865702.

[26] M. Xia, W. Liu, K. Wang, W. Song, C. Chen, and Y. Li, "Non-intrusive load disaggregation based on composite deep long short-term memory network," *Expert Syst. Appl.*, vol. 160, Dec. 2020, Art. no. 113669, doi: 10.1016/j.eswa.2020.113669.

[27] L. Yin, Q. Gao, L. Zhao, and T. Wang, "Expandable deep learning for real-time economic generation dispatch and control of three-state energies based future smart grids," *Energy*, vol. 191, Jan. 2020, Art. no. 116561, doi: 10.1016/j.energy.2019.116561.

[28] M. Figueiredo, A. de Almeida, and B. Ribeiro, "Home electrical signal disaggregation for non-intrusive load monitoring (NILM) systems," *Neurocomputing*, vol. 96, pp. 66–73, Nov. 2012, doi: 10.1016/j.neucom.2011.10.037.

[29] P. Held, S. Mauch, A. Saleh, D. O. Abdeslam, and D. Benyoucef, "Frequency invariant transformation of periodic signals (FIT-PS) for classification in NILM," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5556–5563, Sep. 2019, doi: 10.1109/TSG.2018.2886849.

[30] A. Yang, W. Li, and X. Yang, "Short-term electricity load forecasting based on feature selection and least squares support vector machines," *Knowl.-Based Syst.*, vol. 163, pp. 159–173, Jan. 2019, doi: 10.1016/j.knosys.2018.08.027.

[31] D. Yang, X. Gao, L. Kong, Y. Pang, and B. Zhou, "An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance," *IEEE Trans. Consum. Electron.*, vol. 66, no. 2, pp. 173–182, May 2020, doi: 10.1109/TCE.2020.2977964.

[32] S.-Q. Zhang and Z.-H. Zhou, "Flexible transmitter network," 2020, *arXiv:2004.03839*. [Online]. Available: https://arxiv.org/abs/2004.03839

[33] M. W. Asres, L. Ardito, and E. Patti, "Computational cost analysis and data-driven predictive modeling of cloud-based online NILM algorithm," *IEEE Trans. Cloud Comput.*, early access, Jan. 14, 2021, doi: 10.1109/TCC.2021.3051766.

[34] L. Morán-Fernández, K. Sechidis, V. Bolón-Canedo, A. Alonso-Betanzos, and G. Brown, "Feature selection with limited bit depth mutual information for portable embedded systems," *Knowl.-Based Syst.*, vol. 197, Jun. 2020, Art. no. 105885, doi: 10.1016/j.knosys.2020.105885.

[35] L. Tang, Z. Li, P. Ren, J. Pan, Z. Lu, J. Su, and Z. Meng, "Online and offline based load balance algorithm in cloud computing," *Knowl.-Based Syst.*, vol. 138, pp. 91–104, Dec. 2017, doi: 10.1016/j.knosys.2017.09.040.

[36] G. Zou, Z. Qin, S. Deng, K.-C. Li, Y. Gan, and B. Zhang, "Towards the optimality of service instance selection in mobile edge computing," *Knowl.-Based Syst.*, vol. 217, Apr. 2021, Art. no. 106831, doi: 10.1016/j.knosys.2021.106831.

[37] Z. Zhou, H. Yu, C. Xu, Z. Chang, S. Mumtaz, and J. Rodriguez, "BEGIN: Big data enabled energy-efficient vehicular edge computing," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 82–89, Dec. 2018, doi: 10.1109/MCOM.2018.1700910.

[38] M. Li, P. Si, and Y. Zhang, "Delay-tolerant data traffic to software-defined vehicular networks with mobile edge computing in smart city," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9073–9086, Oct. 2018, doi: 10.1109/TVT.2018.2865211.

CHENXIAO MA was born in Heze, Shandong, China, in 1998. He received the B.S. degree in electrical and its automation from Northeast Electrical Power University, Jilin, China, in 2020. He is currently pursuing the M.S. degree with the College of Electrical Engineering, Guangxi University, Guangxi, China. His major research interest includes artificial intelligence appliance in power systems.

LINFEI YIN was born in Jiujiang, Jiangxi, China, in 1990. He received the B.S. and M.S. degrees in information engineering from Nanchang Hangkong University, Nanchang, China, in 2012 and 2015, respectively, and the Ph.D. degree from the College of Electric Power Engineering, South China University of Technology, Guangzhou, China, in 2018. He is currently an Assistant Professor with the College of Electrical Engineering, Guangxi University, Nanning, Guangxi, China. His major research interest includes artificial intelligence techniques in operation of power systems.