IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

IEEE POWER & ENERGY SOCIETY SECTION

Received December 11, 2020, accepted December 21, 2020, date of publication December 24, 2020,
date of current version January 5, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3047147

A Nonintrusive Load Identification Model Based on Time-Frequency Features Fusion

KEXIN LI[1], BO YIN[1,2], ZEHUA DU[1], AND YUFEI SUN[1]
[1]School of Information Science and Engineering, Ocean University of China, Qingdao 266000, China
[2]Pilot National Laboratory for Marine Science and Technology, Qingdao 266000, China

Corresponding author: Bo Yin (ybfirst@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61972367, and in part by the Key Research and Development Projects of Shandong Province under Grant 2019JMRH0109.

ABSTRACT Nonintrusive load monitoring (NILM) plays a key role in the real-time electricity consumption monitoring of household appliances. However, it is difficult to realize high precision load identification by using a single waveform feature. Therefore, this article proposes a two-stream convolutional neural network based on current time-frequency feature fusion for nonintrusive load identification. First, a time series image coding method for current time-frequency multi-feature fusion is proposed. The method can extract the time domain and frequency domain features of the current timing signal effectively. Then, we present a two-stream neural network combining the gated recurrent unit (GRU) and a two-dimensional convolutional neural network (2D-CNN) to improve the load identification performance. Finally, the experimental results on the PLAID and IDOUC datasets show that the proposed model outperforms the state-of-the-art methods.

INDEX TERMS Nonintrusive load monitoring, appliance classification, feature extraction, signal form.

I. INTRODUCTION

Global energy consumption and greenhouse gas emissions have been increasing in recent years, and household electricity consumption accounts for a large share of total energy consumption [1]. Moreover, it can be demonstrated that appliance-level energy feedback can save 12% of electrical energy in household domestic electricity consumption. Hence, a reasonable reduction in household electricity consumption will improve household electricity efficiency and reduce energy consumption.

Due to the development of technology, smart grids are widely used to monitor and optimize home energy consumption in real-time. As the key technology in acquiring the real-time status of each electrical appliance in an intelligent smart grid, load monitoring technology is mainly divided into intrusive load monitoring (ILM) and nonintrusive load monitoring (NILM). Compared with ILM, NILM has the advantages of easy deployment and low investment costs. The goal of NILM is to obtain the energy consumption of each electrical appliance by decomposing the total energy consumption, and the method is particularly suitable for intelligent monitoring.

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

Since the fluctuation characteristics of electrical signals are not significant, and the changes are weak, the accurate online identification of the load remains challenging. In addition, low power appliances and multistate appliances, which are difficult to identify, occupy a large proportion of common household appliances. In summary, the purpose of this article is to address the multistate and low power appliance recognition problems prevalent in NILM by investigating the conversion of power signals to images.

In NILM, the steady-state characteristics of the active power P and the reactive power Q are commonly used in the early stage [2]. In order to offset the high-frequency information that is cancelled out when calculating active and reactive power, many features such as current waveforms [3], harmonics [4], [5], transient power waveforms [6], and switching transient waveforms [7] are combined on P and Q. In addition to combining P-Q waveform features [2], there are also other methods, such as combining VI trajectories, to convert power signals into images. Although these methods play roles in load identification, it is still necessary to improve the classification of resistive and multistate electrical appliances with similar waveforms. In this article, it is assumed that this coding method can be used to encode a 2D image of the current timing sequence data recorded by smart meters. In which, the algorithm used to generate the current spectrum diagram

used in the data preprocessing phase of this article is based on the concept of the spectrogram in speech recognition, which is the first time that the current spectrum diagram is used for load recognition. This image coding method can make the encoded image contain more relevant information of the original data so as to offset the information loss and help to improve the feature extraction and recognition accuracy of each electrical device.

Hence, a time series image coding method is proposed for current time-frequency multi-feature fusion to obtain the different simultaneous frequency domain features of the current signal. Then, a two-stream convolution network based on the GRU and a 2D-CNN is constructed to obtain a new feature space representation. Finally, different network models are used to classify the appliances. In other words, a load identification model for household electrical appliances with similar power and multistates is built using a new data processing method from the image recognition field.

The structure of the rest of this article is as follows: Chapter 2 provides a summary overview of the work related to household electricity monitoring. Chapter 3 introduces the details of the methods and network frameworks that can improve the effectiveness of electricity monitoring: one encodes the current timing data into two-dimensional images in the data preprocessing stage, and the other is the whole network framework for load identification proposed in this article. Chapter 4 conducts relevant experiments on the PLAID dataset and IDOUC dataset to prove the validity of the proposed methods and models. Finally, Chapter 5 summarizes the findings of this article.

II. RELATED WORK

The NILM method can identify the equipment of aggregate data obtained from a single measurement point [8]. In 1992, hart proposed the preliminary concept of residential load decomposition, explaining how different electrical appliances generate different power signals and showing how the use of equipment is represented by switching events [9]. Although the method performs well, it cannot identify low power devices and devices without discrete power changes [10]. Many researchers have studied this concept and improved the NILM model. H. Najmeddine *et al.* and M.B. Figueiredo *et al.* used current I and voltage V as signatures [11], [12], and they extracted characteristics such as the root mean square (RMS) of the current and peak values. These characteristics are well suited to categorizing kitchen appliances, but the classification results for multistate devices are not as good as those of previous models.

There are also some novel methods that can be used in NILM research, such as the hidden Markov model [13]–[15], machine learning models [16], [17], deep neural networks [18]–[22], etc. Lam *et al.* [23] proposed a classification method for electrical equipment based on the V-I trajectory shape using high-frequency data from the PLAID [24] and WHITED [25] datasets. L.D. Bates [26] used weighted pixelated images of the voltage and current waveform as input data for a deep learning method and then constructed a neural network model consisting of two convolutional layers, which would automatically extract the key features used for the classification of electrical appliances.

Wang and Oates [27], [28] proposed a new method for transforming time series into images that used tiled neural networks to extract features and used a denoising automatic encoder for classification. The results showed that the mean squared error (MSE) obtained by this method could be reduced by up to 48% compared with the original value. L. Kyrkou *et al.* [29] used gramian angular field Matrices (GAF) algorithm to convert power data into images, and then they used VGG16 for migration learning to identify the switching state of a device. The disadvantage of this method is that the migration learning method increases the time complexity.

In NILM, the most common method to improve the classification accuracy is to combine different signal forms such as current-voltage trajectories and active and reactive power. A common drawback of this approach is that it does not take full advantage of all the information in the time and frequency domains of the current data, thus leading to the problem of low load identificatio accuracy. As Pirmin Held [30] mentioned in his paper, until now, no single feature leads to optimal results in the context of NILM because no feature includes all relevant information, and it is difficult to choose the right combination of features. Therefore, with regard to the limitation of using a single device feature for load identification, there are many scholars who have conducted research on this problem, such as [3], [31], [32]. Therefore, the purposes of our research is to propose a feature fusion method that can fully extract the relevant information in the time domain and frequency domain of the current data, so as to improve the identification accuracy of each appliance in load identification.

III. TIME-FREQUENCY FEATURE FUSION MODEL

In order to improve the load identification accuracy for resistive and multistate appliances with similar waveforms in NILM, this article proposes a hybrid model for load identification based on current time-frequency feature fusion. The model includes two main components, namely, feature extraction and load recognition. In the feature extraction section, a signal representation method, which is used to convert the current data into an image that contains as much information as possible, is proposed based on the fusion of the time-frequency domain features of the current timing signal. In the load identification part, a two stream convolutional neural network based on a hybrid load identification model is built to guarantee the overall identification accuracy and improve the training efficiency at the same time. The flowchart of the load identification method is shown in Figure 1.

A. DATA PREPROCESSING

For data preprocessing, the steady-state data of the current are selected as the raw data from which the steady-state

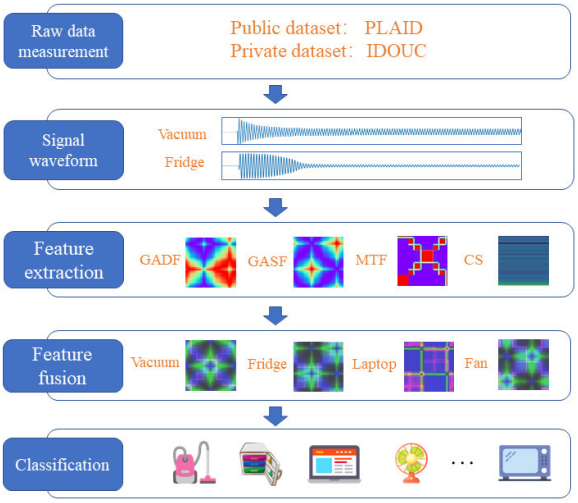


FIGURE 1. The flowchart of the load identification method.

characteristics of the current are extracted. A steady-state characteristic is a characteristic that an appliance exhibits when it is in a stable operating state. Steady-state data are easier to collect and detect than transient data.

In this part, the steady-state data of the current are selected as the raw data to extract the steady-state characteristics, which are 2D images obtained by encoding the time series of the raw data. To achieve this, a time-frequency domain feature fusion model is proposed, as shown in Figure 2. It includes three encoding frameworks for encoding time series into images, namely, the gramian angular summation field (GASF), gramian angular difference field (GADF), and markov transition field (MTF) algorithms, which can preserve the time dependence of current timing data and preserve the static and dynamic information in the time domain of the current signal [27]. Additionally, a current spectrogram (CS) is also contained in the fusion model to represent the spectral information of the current timing data. The introduction of the CS enhances the frequency domain information in the image, and this is the first time that a current frequency domain analysis map is used for electricity use identification.

1) GAF AND MTF ALGORITHM

The GAF method constructs an image by representing the time series in polar coordinates, which requires scaling the used time series to an interval of [-1,1] or [0,1]. After the recalibrated time series is converted to the polar coordinate system, the GASF and GADF are defined by considering the trigonometric sum and differences between the points:

$$\begin{aligned} GASF &= [cos(\phi_i + \phi_j)] \\ &= \tilde{X}' \cdot \tilde{X} - \sqrt{I - \tilde{X}^2}' \cdot \sqrt{I - \tilde{X}^2}. \end{aligned} \tag{1}$$

$$\begin{aligned} GADF &= [cos(\phi_i - \phi_j)] \\ &= \sqrt{I - \tilde{X}^2}' \cdot \tilde{X}' \cdot \tilde{X} - \sqrt{I - \tilde{X}^2}. \end{aligned} \tag{2}$$

where $\widetilde{X}$ represent the rescaled time series in polar coordinate, and ϕ represent the value as the angular cosine. The advantage of GAF encoding is that it constructs a bijection map between a one-dimensional time series and a two-dimensional space so that no information is lost and time dependence can be maintained through r coordinates. The diagonal of the Gram-like matrix consists of the original values of the scaled time series, which means that the time series can be approximately reconstructed according to the high-level features learned by the deep neural network.

The MTF algorithm can encode dynamic information by continuously representing the Markov transformation probability and saving information in the time domain. w_{ij} is given by the frequency with which a point in the quantile q_j is followed by a point in the quantile q_i. In order to overcome the loss of information caused by removing time dependence from matrix W, the MTF is defined as follows:

$$MTF = \begin{pmatrix} w_{ij}|_{x_1 \in q_i, x_1 \in q_j} & \cdots & w_{ij}|_{x_1 \in q_i, x_n \in q_j} \\ \vdots & \ddots & \vdots \\ w_{ij}|_{x_n \in q_i, x_1 \in q_j} & \cdots & w_{ij}|_{x_n \in q_i, x_n \in q_j} \end{pmatrix} \tag{3}$$

2) CURRENT SPECTRUM DIAGRAM

It is known that different electrical circuits will lead to different harmonics. In other words, the harmonic data contain the unique characteristics of different electrical appliances, and it is possible to recognize electrical appliances via the harmonic data of the load voltage or current. Hence, the current spectrum diagram, which is produced by using the Fourier transform or wavelet transform, is proposed according to the concept of the spectrogram in speech recognition, as shown in Figure 3. It can be produced by using the Fourier transform. Mathematically speaking, for a current signal $x(t)$, first we divide the frame into $x(m, n)$, (n is the frame length and m is the number of frames), then we perform the FFT transformation to get $X(m, n)$, after that we make a periodogram $Y(m, n)$, then we transform m into a scale M according to time and n into a scale N according to frequency, then we draw into a two-dimensional graph which is the current spectrum diagram. It is a type of 2D spectrum diagram that contains three-dimensional information and represents the graph of the current spectrum changing with the time and frequency. As shown, the darker the color, the greater the current intensity. The current spectrum will differ because the electrical equipment creates different depths that form different grains; that is, the current spectra of different electrical equipment are different. Therefore, this type of diagram can be used to identify different electrical appliances.

3) CONSTRUCTION OF THE FEATURE FUSION MODEL

It is hard to explain the collinear and potential states of electrical appliances by using univariate time series data. In this article, a new time-frequency domain feature fusion model based on a 2D image coding method is proposed to retain the time information of the original data as much as possible. In this model, the GAF and MTF are used to encode the static and dynamic current series information in the time domain. In addition, the current spectrum diagram is used to describe

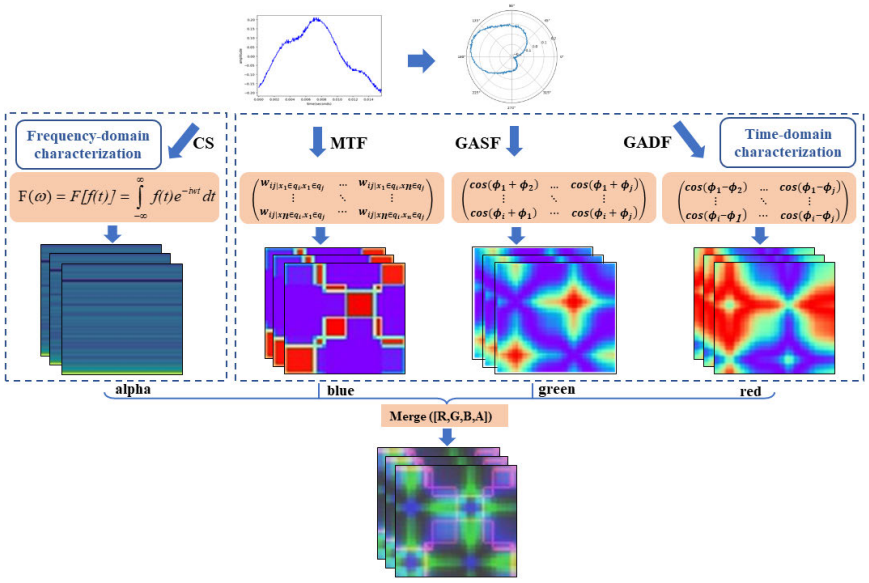


FIGURE 2. Time-frequency domain feature fusion model.

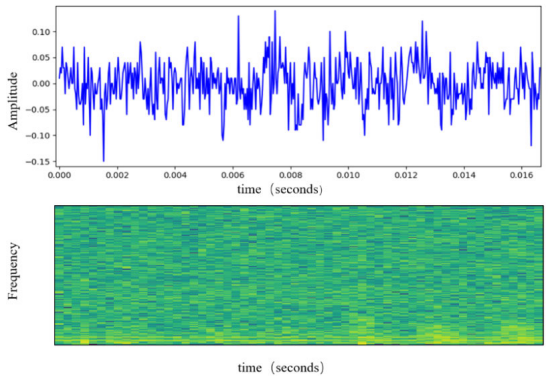


FIGURE 3. Current spectrum diagram.

the frequency domain characteristics of the current timing data. In this way, a four-channel image will be generated by combining all the time-frequency domain features extracted from the GASF, GADF, MTF and CS. Therefore, the static information, dynamic information and spectral information of the original time series are all retained and considered at the same time. The time series image coding in the NILM (TSICN) algorithm where the time-frequency domain features fusion model of the current signal constructed in this article is shown in Algorithm 1.

Algorithm 1 Time Series Image Coding in NILM

Input: D: Dataset; P_Hz: Power frequency; D_Hz: Dataset sampling frequency; t: Samping time; I_value: Current data for each sample in the dataset; N: the number of categories in the dataset;

Output: D': Dataset of single period current after two-dimensional encoded image fusion;

Calculate of the number of sampling points of single cycle current current: $Sp\_Num \leftarrow \frac{1}{P\_Hz} \times D\_Hz$;
Calculate the number of current cycles contained in the sampling time: $Cycle\_Num \leftarrow P\_Hz \times t$;
for $I\_value[0]$ to $I\_value[N-1]$ **do**
    Every Sp_Num pieces of data are divided into groups
    for *the data in each group* **do**
        a) Generate four two-dimensional images, respectively using GADF, GASF, MTF,CS algorithm
        b) Gets the grayscale values of the above four images, save them as R, G, B and A and convert A to be between 0 and 1.
        c) Merge four images into one image Merge$([R, G, B, A])$.
    end for
end for

B. TWO-STREAM CONVOLUTIONAL NEURAL NETWORK

In order to improve the load identification performance and the convergence speed of network models, a two-stream convolution network based on the GRU and the 2D-CNN is constructed (as shown in Figure 4)

In the GRU section, to combine the speed and lightness of a one-dimensional convolutional network and the sequential sensitivity of the RNN, we adopted the following network structure. The input data are the raw current signal, and the single cycle current data are selected as samples. We use a 1D-CNN in front of the GRU as a preprocessing step, and the convolutional neural network converts long input sequences into shorter sequences with an advanced feature composition. The extracted features then compose these sequences that become inputs to the GRU in the network. This ensures that the training speed of the model is improved. A 3-layer convolutional kernel with a size of 7 and a step length of 1 is

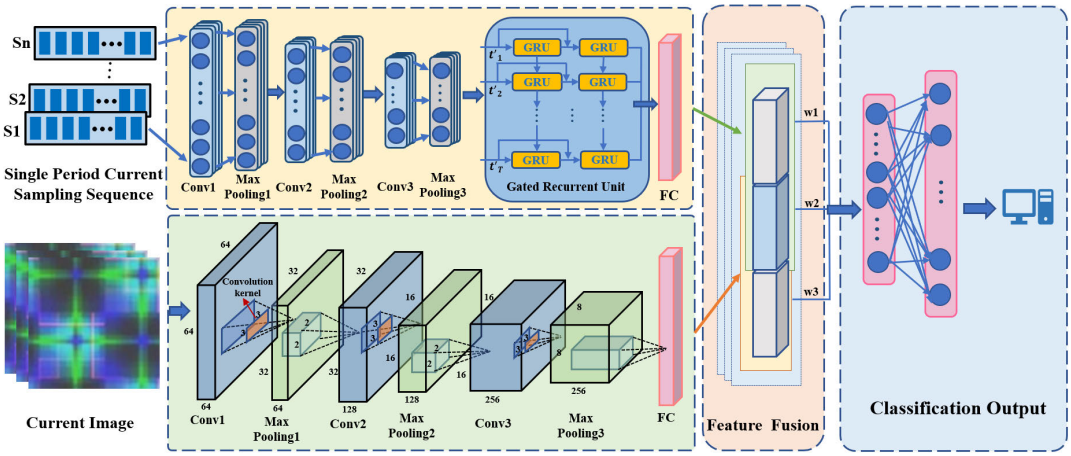


FIGURE 4. Overall network structure diagram.

used to preprocess the input data. The data are then entered into the GRU layer for further feature extraction. Furthermore, in order to solve the problem of low load recognition accuracy caused by extracting only a single waveform feature, we designed the following network structure. The input data are two-dimensional images with a size of 64×64 converted from the original data by the TSICN algorithm. When processing images, convolutional core filters in CNN networks are often used to learn features with different levels of abstraction. Because the CNN has allows weight sharing, network parameters can be greatly reduced. In addition, its local perception structure can improve the ability to learn the local details of the load. Therefore, we use a three-layer, two-dimensional convolution to extract features. We use a convolution kernel of size 3 and a step size of 1, and the number of convolution kernel increases in turn. In addition, batch normalization (BN) is used in front of the fully connected layer of each CNN stream to reduce the number of parameters, which are not marked in the figure.

When current sequences and image data are utilized simultaneously, they are often simply treated as undifferentiated multichannel data, and after feature extraction with different networks, the last fully connected layer is fused. These methods do not take into account the differences in features between the different modalities when feature fusion is performed. Based on this, this article adopts the ideas in [33] when designing a fusion framework for two-stream convolutional neural network models. Based on the last extracted features, the consistency or shared features between the two modes are considered. The features of each modality should contain the independent features of its own modality and the features of the related parts of both modalities. Since the influences of different modalities on the load identification result are not consistent and the importance of the independent features and the related features is also not consistent, the two modal features should be considered. Therefore, by determining the weights of the different features, more distinguishable and robust fusion features can be generated.

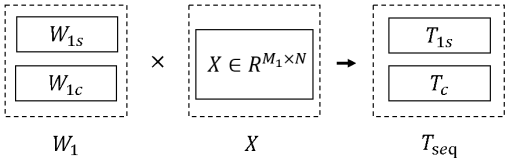


FIGURE 5. Illustration of the transformation matrix.

In order to achieve this goal, we analyze the extracted features for different networks and classify the final extracted features of the two networks into independent features and shared features after the feature transformation matrix. As shown in Figure 5, $X = [x_1, x_2, \ldots x_N] \in R^{M_1 \times N}$ represents the M-dimensional output of the last fully connected layer of the network in a batch of N current sequences. $W_1 = [W_{1s}; W_{1c}]$ denotes the sum of the transformation matrices corresponding to the independent and correlated features, respectively, when the current sequence is used as input, by which the changes in this matrix can be used to generate the independent and correlated features under the current network. where $W_{1s} \in R^{K_{1s} \times M_1}$ denotes the transformation matrix that generates the relevant features of the current series, $W_{1c} \in R^{K_c \times M_1}$ denotes the transformation matrix that generates the independent features of the current series, and $M_1' = K_{1s} + K_c$. $T_{1s} \in R^{K_{1s} \times M_1}$ and $T_c \in R^{K_c \times M_1}$ represent the independent and correlation features generated when the Input is a current sequence, respectively, and $T_{seq} = [T_{1s}; T_c]$. The representation of the input current image is the same as described above. Therefore, we can find the transformation matrices W_1 and W_2 to obtain T_{seq} and T_{Image}, and finally splice the two features to obtain $T = [T_c; T_{1s}; T_{2s}]$.

In addition, when determining the weights of different modal features, it is possible to determine which modal data are more important without having to know in advance which modal data are more important simply by using the framework. Performing cyclic iterative learning, as shown in Figure 6, can accomplish adaptive weight control.

In the figure, $W^{(c)T}$, $W^{(1s)T}$ and $W^{(2s)T}$ correspond to the relevant partial feature weights, independent feature weights

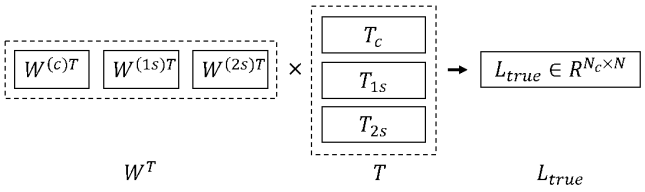


FIGURE 6. Illustration of the regression coefficient matrix.

of current sequence, independent feature weights of current image, and independent feature weights of current image, respectively. weights. $W^T = [W^{(c)T}; W^{(1s)T}; W^{(2s)T}]$ is the regression coefficient matrix, T is the fusion of the final output features of the two-stream network, and L_{true} is the true category label of the appliance. By designing the features of the two-stream fully connected layer output, the fused features and corresponding labels are obtained. Finally, the fused feature data and labels are input to the multilayer perceptron for classification, thereby identifying the appliance category and completing the load identification.

IV. THE REALIZATION OF THE IMPROVED METHOD

A. DATA PREPROCESSING

1) INTRODUCTION TO DATASETS

Two smart electricity datasets, PLAID and IDOUC, were used to validate the proposed method. The PLAID dataset is a public dataset based on actual measurements of 11 electrical appliances in 56 U.S. homes, and it includes current and voltage measurements at a sampling rate of 30 kHz. Each device type is represented by dozens of different examples of different makes and models. For each device, 3 to 6 measurements are collected per state transition. There are 1074 instances in PLAID1 and 719 instances in PLAID2. The data from PLAID1 is used in this paper. The types of appliances in the dataset are Fan, Hairdryer, Heater, Microwave, Vacuum, Air Conditioner (AC), Incandescent Light Bulb (ILB), Compact Fluorescent Lamp (CFL),Washing Machine (WM), Fridge and Laptop.

IDOUC is the smart electricity dataset of Ocean University of China. The data collected in the dataset consisted of current and voltage data from 103 different appliance types at a sampling rate of 4kHz. For each appliance, 100 data samples containing steady-state and transient data were collected separately. The data in the dataset are obtained by simulating the household electricity environment in the laboratory. The measured loads are all the common residential loads, and then sensors are used to collect the electricity data of the main trunk road of the residential power supply. The model of the data acquisition device is established as shown in Figure 7.

2) MULTI-FEATURE EXTRACTION AND FUSION

Figure 8 shows the typical current and voltage waveforms in the PLAID and IDOUC datasets, respectively. Both transient and steady-state of current waveform are contained. In this study, the current steady-state data of 60 cycles is intercepted as raw data, shown in Figure 8. Steady-state data is easier to collect and detect than transient data. For each device current measurement in this study, the current steady-state data for the 60 cycles before the appliance is turned off is intercepted as raw data. Figure 9 shows the images obtained by TSICN algorithm coding with the single-period current data of each device. It can be seen from the images of each device that their images differ significantly, so different loads can be distinguished according to different features of the images.

B. EXPERIMENTAL AND RESULTS

In order to make a comprehensive evaluation of the proposed model, not only the performance of the whole model, but also the performance of the key part of the model is studied. First, in order to compare the effectiveness of time series coding, the TSICN algorithm is compared with GAF, MTF, CS and other coding methods. meanwhile, it is also compared with the widely used VI trajectory characteristics. According to the operation of [34], this study selects and compares the VI trajectory after gridding. Then, we compare the performance of the images encoded by the TSICN algorithm and the original data as inputs of different neural networks. Finally, in order to evaluate the overall performance of the proposed model, the results of the proposed model are compared with the results of the existing model. In the existing household appliance classification model, this article chooses to compare with the model in [35], because the above model is also built on all instances of PLAID data set and adopted the method of cross-validation.

This chapter describes the details of the operation and network structure of converting data into images. Here, the experiment is divided into three parts, focusing on the analysis of different image coding methods, comparison of the performance of different neural networks and the experimental results of the proposed network model.

1) EVALUATION INDEX

For classification problems, there are two methods to calculate the accuracy. Calculating the accuracy of category predictions without differentiating categories is called micro-averaging. Another method is to calculate the accuracy of the predictions for each category and then calculate the accuracy of the predictions through arithmetic averages, called macro averages. In this article, the macro average is selected as the evaluation index, and the obtained equation is as follows [36]:

$$P_{macro} = \frac{1}{n}\sum_{i=1}^{n} P_i \tag{4}$$

$$R_{macro} = \frac{1}{n}\sum_{i=1}^{n} R_i \tag{5}$$

$$F_{macro} = \frac{2 \times P_{macro} \times R_{macro}}{P_{macro} + R_{macro}} \tag{6}$$

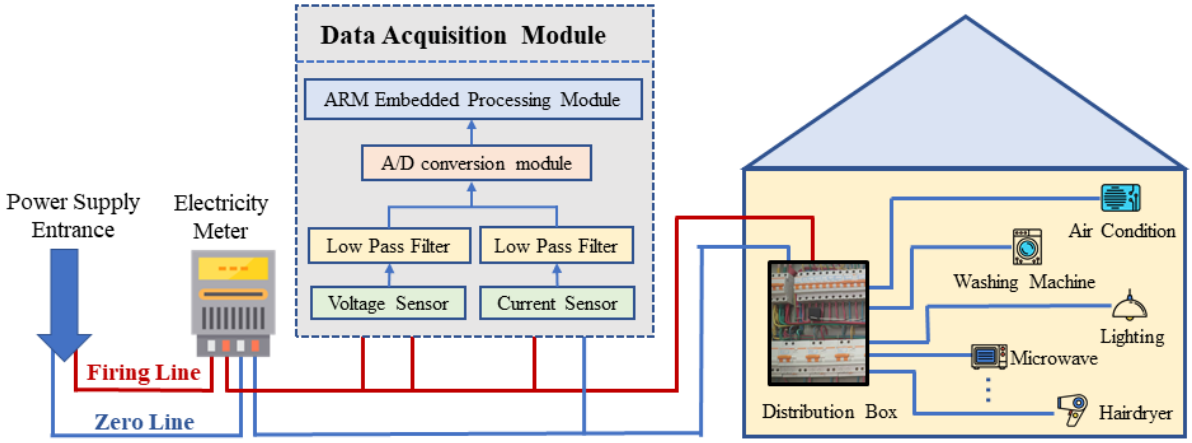


FIGURE 7. Data acquisition device model.

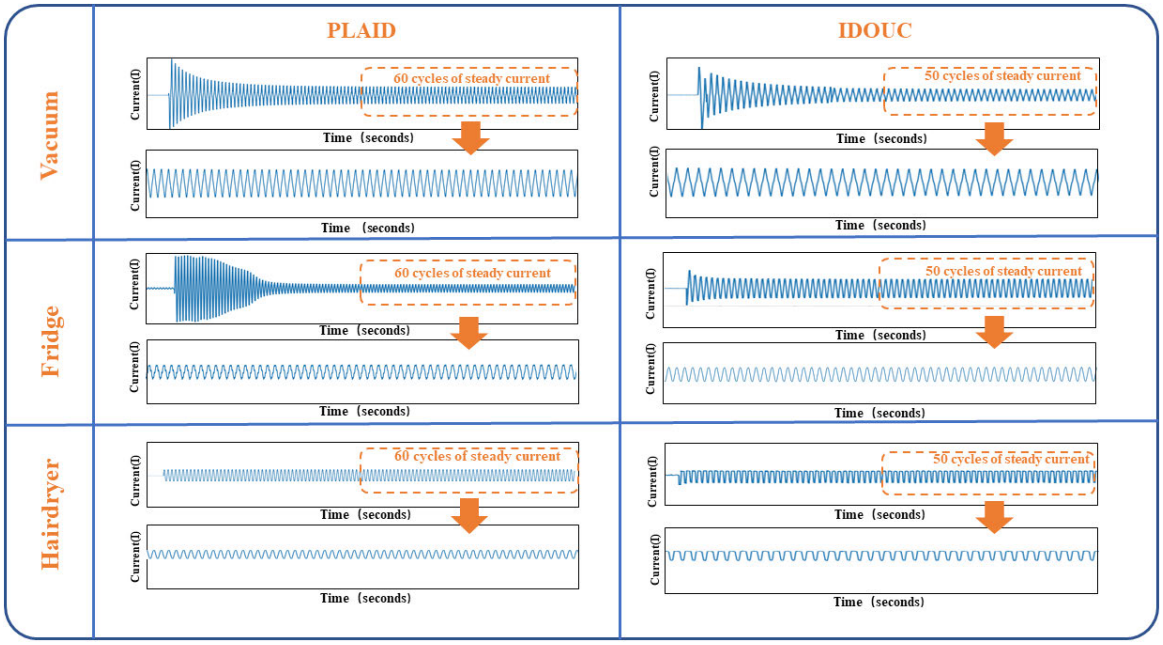


FIGURE 8. Current waveforms from a vacuum in the extended PLAID: Current waveform of vacuum over 2s and 60 cycles of steady current of vacuum.

where $P_i = \frac{TP_i}{TP_i+FP_i}$, $R_i = \frac{TP_i}{TP_i+FN_i}$, TP_i represents the number of samples that predict the positive class as the positive class, and FP_i represents the number of samples that predict the negative class as the positive class.

2) COMPARISON OF DIFFERENT FEATURES

In order to prove the feasibility of the image coding method proposed in this article, a comparative experiment is designed, in which the current data are coded by different methods respectively for comparison. As can be seen from Table 1, the image encoded by GADF algorithm has the best classification effect among the single coding methods. In terms of overall comparison, the classification effect of images encoded by TSICN algorithm after fusion is better than any single coding method.

TABLE 1. The accuracy results of applying different coding methods to the current data in the PLAID dataset.

Model	GADF	GASF	MTF	CS	VI	TSICN
2D-CNN	97.68%	97.85%	83.47%	92.34%	97.26%	98.27%
2D-CNN-BN	98.25%	98.16%	84.55%	94.47%	97.44%	98.51%

3) COMPARISON OF DIFFERENT NEURAL NETWORK

In order to prove that the use of coded images is more beneficial to improve the recognition accuracy than the use of original data, as well as the performance and benefits of GRU over one-dimensional convolution neural network (1D-CNN) for the original current data on PLAID dataset, a comparison experiment is designed in this chapter. The GRU network model used is shown in the top half of Figure 4.

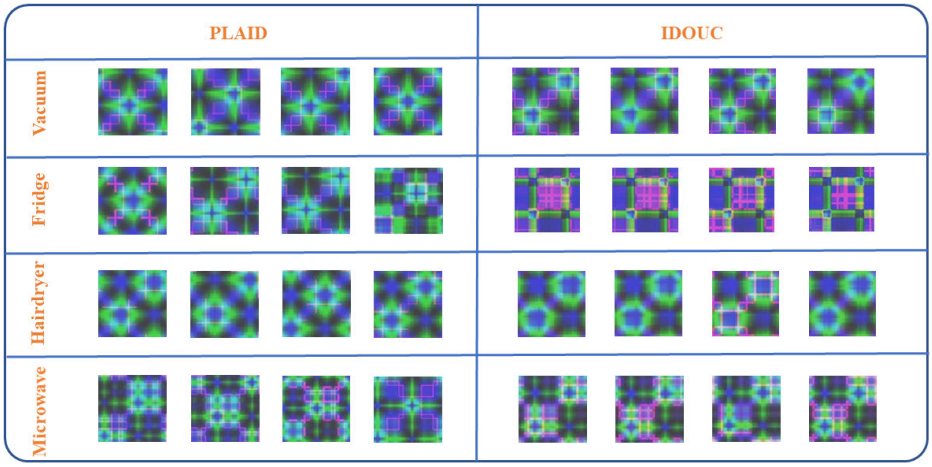
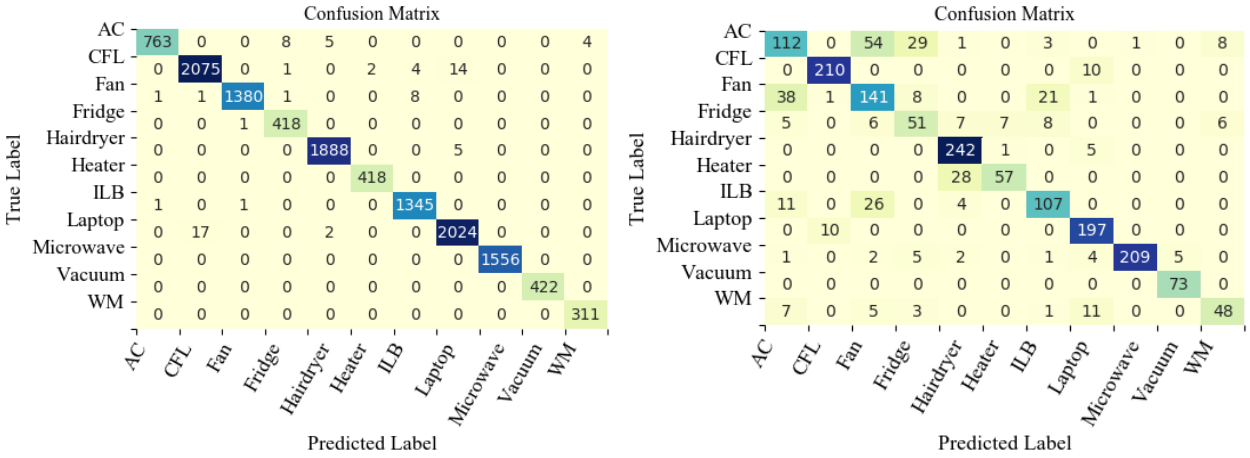


FIGURE 9. Image obtained by merging single-cycle current data of four appliances in the PLAID dataset and the IDOUC dataset.



(a) Confusion Matrix of the proposed model

(b) Confusion Matrix of the state of the art model

FIGURE 10. The confusion matrix of the proposed model and the state of the art model on 11 appliances.

TABLE 2. Comparison of accuracy results of different neural networks and different inputs.

Input	Current		Image
Model	1D-CNN-BN	GRU	2D-CNN-BN
Accuracy	90.22%	95.87%	98.51%

It can be seen from Table 2 that in terms of the performance and benefit of model training, the classification accuracy rate of GRU is 95.87%, which is significantly better than 1DCNN. Moreover, 1DCNN has a faster convergence rate during training. Although 2DCNN had the highest accuracy of 98.51%, it took about 18 times longer to train than the GRU model.

4) PERFORMANCE OF THE OVERALL NETWORK STRUCTURE

In this section, the proposed network structure is a two-stream model combined with 2D-CNN and GRU, which is compared with two single-input network models.

For 11 appliances in the PLAID dataset, the steady-state part of the original dataset was intercepted to draw a two-dimensional image of the current. Figure 10 shows the confusion matrix of the proposed model. Based on the confusion matrix, the classification precision, recall rate and F1 score of each device are calculated. It can be seen from Figure 10 and Table 3 that:

At first, the overall performance of the proposed model is better than that of the existing model. For example, the accuracy and F1-macro of the proposed model are 0.9949 and 0.9939 respectively, while the accuracy and F1-macro of the State of the art model are 0.8070 and 0.7871 respectively. Furthermore, for each device, the proposed model has better performance than the existing model. As shown in Table 3, compared with the existing model, the proposed model has a higher F1 value for 11 devices in the dataset, especially for 6 devices, namely AC, Fan, Fridge, Heater, ILB and WM. Finally, for air conditioning, fridge, washing machine, laptop, heater such as multi-state electrical appliances and similar electrical waveform classification performance has

TABLE 3. The performance of the proposed model and the state of the art model on 11 appliances.

Model	Appliance	Precision	Recall	F1	F1-marco	Acc
Proposed Model	AC	0.9782	0.9974	0.9877	0.9939	0.9949
	CFL	0.9900	0.9914	0.9907		
	Fan	0.9921	0.9986	0.9953		
	Fridge	0.9976	0.9766	0.9870		
	Hairdryer	0.9974	0.9963	0.9968		
	Heater	1.0000	0.9952	0.9976		
	ILB	0.9985	0.9896	0.9940		
	Laptop	0.9902	0.9912	0.9907		
	Microwave	1.0000	1.0000	1.0000		
	Vacuum	1.0000	1.0000	1.0000		
	WM	1.0000	0.9871	0.9935		
State of the art model	AC	0.6437	0.5385	0.5864	0.7871	0.8070
	CFL	0.9502	0.9545	0.9524		
	Fan	0.6026	0.6714	0.6351		
	Fridge	0.5313	0.5667	0.5484		
	Hairdryer	0.8521	0.9758	0.9098		
	Heater	0.8769	0.6706	0.7600		
	ILB	0.7589	0.7230	0.7405		
	Laptop	0.8640	0.9517	0.9057		
	Microwave	0.9952	0.9127	0.9522		
	Vacuum	0.9359	1.0000	0.9669		
	WM	0.7742	0.6400	0.7007		

been significantly improved. For example, the F1-score of air conditioners, fridge, washing machines, laptops and heaters increased by 40.62%, 44.43%, 29.47%, 8.5% and 23.81%, respectively. In addition, it can also be seen from Table 3 that the accuracy of fridge and washing machine is low, which is due to the complicated working process of refrigerator and washing machine, which will produce various waveforms, leading to the confusion with the waveforms of other devices.

For each of the 21 appliances in the IDOUC dataset, 10,000 converted 2D-images are eventually available for each device. Through the experiment, it can be concluded that the evaluation index F_{macro}=99.12%. Figure 11 (a) shows the training results of the proposed model represented by the confusion matrix, in which the minimum value comes from the fridge class, and the Confusion Matrix of the state of art model are shown in Figure 11 (b). The Y-axis is the real label of the appliance, and the X-axis is the label predicted by the model. From this diagram you can see the probability that the appliance is misidentified as another device. It is found that the power waveform generated when the compressor works in the fridge will be confused with the freezer and mini-fridge, and when the temperature in the fridge reaches the set value, the fridge will enter the standby state, the power waveform generated at this time will be confused with the low-power electrical appliances, resulting in a decrease in the recognition rate.

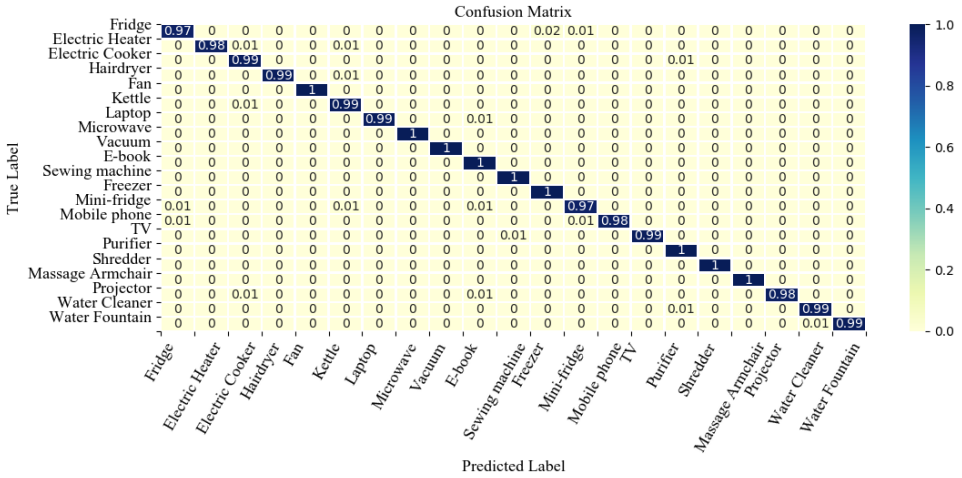
Figure 12 is a boxplot of the accuracy obtained after multiple experiments using different neural network models in the fused image, where the solid line in the middle represents the median accuracy obtained from the test data. As can be seen from Figure 12 and Table 4, the network model

TABLE 4. The accuracy of classification on different datasets using different network models.

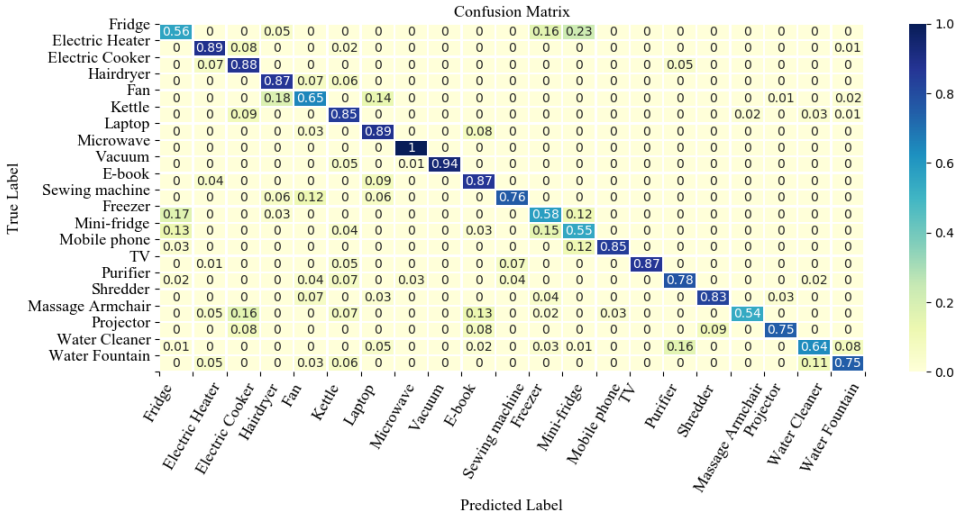
Model	Dataset	Vgg16	2D-CNN-BN	Fusion model (Add)	This paper
GAF-MTF-CS	PLAID	96.24%	98.51%	98.92%	99.49%
	IDOUC	97.12%	97.82%	97.89%	99.78%

proposed in this article has a better classification effect than other network models for TSICN-encoded images. At the same time, it can be seen from Figure 13 that during the model training, the convergence speed of the network model is faster than that of 2D-CNN, and the accuracy is improved compared with 2D-CNN, and the training time of a round is about 0.2 times that of 2D-CNN. Therefore, we can see from the experimental results that the addition of the GRU section can speed up the convergence of the overall model, thus compensating for the shortcomings in the convergence speed of the 2D-CNN model.

In order to make a comprehensive evaluation of the proposed model, the validity of the model was verified on the PLAID dataset and the IDOUC dataset, and several comparative experiments were conducted. The comparison results show that: At first, the F1 score of the model on the PLAID dataset is 0.9939, and the overall performance of the model is better than that of the existing model. Then, compared with the existing model, this model has significantly improved the classification performance of multi-state appliances, such as refrigerators, washing machines, laptops and heaters, and appliances with similar waveforms. Furthermore, compared



(a) Confusion Matrix of the proposed model



(b) Confusion Matrix of the state of the art model

FIGURE 11. The confusion matrix of the proposed model and the state of the art model on the IDOUC dataset.

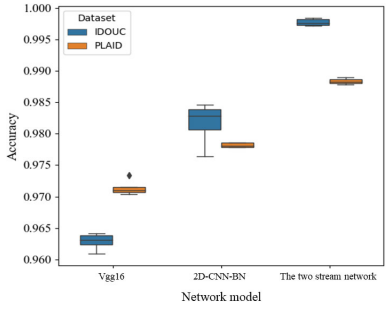


FIGURE 12. Accuracy of the merged image using Vgg16,2D-CNN-BN and the two stream network.

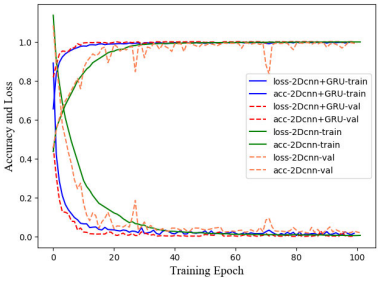


FIGURE 13. Loss and accuracy of the merged image using 2D-CNN-BN and the two stream network.

with the widely used current and voltage characteristics, the time-frequency domain features integrated by TSICN method provide better overall performance in distinguishing electrical equipment. Moreover,we compare the overall model, by comparing it with the traditional feature fusion model, which is achieved at the last layer of the network model using the addition fusion of feature vectors at the full connection layer, and it can demonstrate that the feature fusion method used in this article has a higher identification rate than other traditional fusion approaches and is more applicable to existing network structures.

Therefore, it can be concluded from the above experimental results that this method can effectively identify multi-state electrical appliances and similar waveform electrical appliances, and performs well in detecting devices that are difficult to identify by other models.

V. CONCULSION

To address the limitation of using a single appliance feature for load identification, this article proposes a feature extraction method based on the time-frequency feature fusion of current time series signals. To improve the recognition accuracy of multistate appliances and similar waveform appliances that occur in NILM, we designed a two-stream convolution network combining with the GRU and 2D-CNN. The effectiveness of the proposed model is verified on the PLAID dataset and IDOUC dataset. The identification accuracies of the model trained on the PLAID and IDOUC datasets of the acquired model are 99.49% and 99.78%, respectively. Several comparison results prove that the time-frequency feature fusion model constructed by this method can convert one-dimensional time series into two-dimensional images and retain all the timefrequency domain information of the original signal. Furthermore, it can be verified that the two-stream neural network model has a good classification effect and a faster model training rate while ensuring the load identification accuracy. This article lays the foundation for the further separation of power consumption fluctuation signals in the intelligent monitoring of household power consumption. In the future, we plan to apply the method to event detection and multilabel classification of complex scenarios in NILM.

REFERENCES

[1] D. Wang, Y. Song, and Y. Zhu, "Information platform of smart grid based on cloud computing," *Autom. Electr. Power Syst.*, vol. 34, no. 22, pp. 7–12, Jul. 2010.

[2] D. Weisshaar, P. Held, S. Mauch, and D. Benyoucef, "Device classification for NILM using FIT-PS compared with standard signal forms," in *Proc. Int. IEEE Conf. Workshop Óbuda Electr. Power Eng. (CANDO-EPE)*, Nov. 2018, pp. 1–6, doi: 10.1109/CANDO-EPE.2018.8601150.

[3] L. Du, J. A. Restrepo, Y. Yang, R. G. Harley, and T. G. Habetler, "Nonintrusive, self-organizing, and probabilistic classification and identification of plugged-in electric loads," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1371–1380, Sep. 2013, doi: 10.1109/TSG.2013.2263231.

[4] S. Semwal, G. Shah, and R. S. Prasad, "Identification residential appliance using NIALM," in *Proc. IEEE Int. Conf. Power Electron., Drives Energy Syst. (PEDES)*, Dec. 2014, pp. 1–6, doi: 10.1109/PEDES.2014.7041965.

[5] J. Lei, S. Luo, and J. Li, "An approach of household power appliance monitoring based on machine learning," in *Proc. Int. Conf. Intell. Comput. Technol. Automat.*, 2012, pp. 577–580.

[6] S. Rahimi, A. D. C. Chan, and R. A. Goubran, "Nonintrusive load monitoring of electrical devices in health smart homes," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2012, pp. 2313–2316, doi: 10.1109/I2MTC.2012.6229453.

[7] M. Azaza and F. Wallin, "Supervised household's loads pattern recognition," in *Proc. IEEE Electr. Power Energy Conf. (EPEC)*, Oct. 2016, pp. 1–5, doi: 10.1109/EPEC.2016.7771718.

[8] A. Zoha, A. Gluhak, M. Imran, and S. Rajasegarar, "Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey," *Sensors*, vol. 12, no. 12, pp. 16838–16866, Dec. 2012, doi: 10.3390/s121216838.

[9] G. W. Hart, "Nonintrusive appliance load monitoring," *Proc. IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992, doi: 10.1109/5.192069.

[10] L. Du, Y. Yang, D. He, R. G. Harley, and T. G. Habetler, "Feature extraction for load identification using long-term operating waveforms," *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp. 819–826, Mar. 2015, doi: 10.1109/TSG.2014.2373314.

[11] H. Najmeddine, K. El K. Drissi, C. Pasquier, C. Faure, K. Kerroum, A. Diop, T. Jouannet, and M. Michou, "State of art on load monitoring methods," in *Proc. IEEE 2nd Int. Power Energy Conf.*, Johor Bahru, Malaysia, Dec. 2008, pp. 1256–1258, doi: 10.1109/PECON.2008.4762669.

[12] M. B. Figueiredo, A. de Almeida, and B. Ribeiro, "An experimental study on electrical signature identification of non-intrusive load monitoring (NILM) systems," in *Adaptive and Natural Computing Algorithms*. Berlin, Germany: Springer, 2011, pp. 31–40, doi: 10.1007/978-3-642-20267-4_4.

[13] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial HMMs with application to energy disaggregation," in *Proc. 15th Int. Conf. Artif. Intell. Statist. (AISTATS)*, Apr. 2012, pp. 1472–1482.

[14] F. Paradiso, F. Paganelli, D. Giuli, and S. Capobianco, "Context-based energy disaggregation in smart homes," *Future Internet*, vol. 8, no. 1, p. 4, Jan. 2016, doi: 10.3390/fi8010004.

[15] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *Proc. Natl. Conf. Artif. Intell.*, vol. 1, Jan. 2012, pp. 1–7.

[16] J. Kim, T.-T.-H. Le, and H. Kim, "Nonintrusive load monitoring based on advanced deep learning and novel signature," *Comput. Intell. Neurosci.*, vol. 2017, pp. 1–22, Oct. 2017, doi: 10.1155/2017/4216281.

[17] P. R. Z. Taveira, C. H. V. De Moraes, and G. Lambert-Torres, "Non-intrusive identification of loads by random forest and fireworks optimization," *IEEE Access*, vol. 8, pp. 75060–75072, 2020, doi: 10.1109/ACCESS.2020.2988366.

[18] T.-T.-H. Le, H. Kang, and H. Kim, "Household appliance classification using lower odd-numbered harmonics and the bagging decision tree," *IEEE Access*, vol. 8, pp. 55937–55952, 2020, doi: 10.1109/ACCESS.2020.2981969.

[19] J. Kelly and W. Knottenbelt, "Neural NILM: Deep neural networks applied to energy disaggregation," in *Proc. 2nd ACM Int. Conf. Embedded Syst. Energy-Efficient Built Environ. (BuildSys)*, Nov. 2015, pp. 55–64, doi: 10.1145/2821650.2821672.

[20] I. H. Çavdar and V. Faryad, "New design of a supervised energy disaggregation model based on the deep neural network for a smart grid," *Energies*, vol. 12, no. 7, p. 1217, Mar. 2019, doi: 10.3390/en12071217.

[21] P. Schirmer and I. Mporas, "Statistical and electrical features evaluation for electrical appliances energy disaggregation," *Sustainability*, vol. 11, no. 11, p. 3222, Jun. 2019, doi: 10.3390/su11113222.

[22] L. R. Morais and A. R. G. Castro, "Competitive autoassociative neural networks for electrical appliance identification for non-intrusive load monitoring," *IEEE Access*, vol. 7, pp. 111746–111755, 2019, doi: 10.1109/ACCESS.2019.2934019.

[23] H. Lam, G. Fung, and W. Lee, "A novel method to construct taxonomy electrical appliances based on load signaturesof," *IEEE Trans. Consum. Electron.*, vol. 53, no. 2, pp. 653–660, May 2007, doi: 10.1109/TCE.2007.381742.

[24] J. Gao, S. Giri, E. C. Kara, and M. Bergés, "PLAID: A public dataset of high-resoultion electrical appliance measurements for load identification research," in *Proc. 1st ACM Conf. Embedded Syst. Energy-Efficient Buildings (BuildSys)*, New York, NY, USA, 2014, pp. 198–199, doi: 10.1145/2674061.2675032.

[25] M. Kahl, A. U. Haq, T. Kriechbaumer, and H.-A. Jacobsen, "Whited—A worldwide household and industry transient energy data set," in *Proc. 3rd Int. Workshop Non-Intrusive Load Monitor.*, May 2016, pp. 14–15.

[26] L. De Baets, T. Dhaene, D. Deschrijver, C. Develder, and M. Berges, "VI-based appliance classification using aggregated power consumption data," in *Proc. IEEE Int. Conf. Smart Comput. (SMARTCOMP)*, Jun. 2018, pp. 179–186, doi: 10.1109/SMARTCOMP.2018.00089.

[27] Z. Wang and T. Oates, "Encoding time series as images for visual inspection and classification using tiled convolutional neural networks," in *Proc. Workshops 29th AAAI Conf. Artif. Intell.*, vol. WS-15-14, Jan. 2015, pp. 40–46.

[28] Z. Wang and T. Oates, "Imaging time-series to improve classification and imputation," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, May 2015, pp. 1–7.

[29] L. Kyrkou, C. Nalmpantis, and D. Vrakas, "Imaging time-series for NILM," *Eng. Appl. Neural Netw.*, 2019, pp. 188–196.

[30] P. Held, F. Laasch, D. O. Abdeslam, and D. Benyoucef, "Frequency invariant transformation of periodic signals (FIT-PS) for signal representation in NILM," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2016, pp. 5149–5154, doi: 10.1109/IECON.2016.7793617.
[31] L. Du, D. He, R. G. Harley, and T. G. Habetler, "Electric load classification by binary voltage–current trajectory mapping," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 358–365, Jan. 2016, doi: 10.1109/TSG.2015.2442225.
[32] N. Sadeghianpourhamami, J. Ruyssinck, D. Deschrijver, T. Dhaene, and C. Develder, "Comprehensive feature selection for appliance classification in NILM," *Energy Buildings*, vol. 151, pp. 98–106, Sep. 2017, doi: 10.1016/j.enbuild.2017.06.042.
[33] A. Wang, J. Cai, J. Lu, and T.-J. Cham, "MMSS: Multi-modal sharable and specific feature learning for RGB-D object recognition," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1125–1133, doi: 10.1109/ICCV.2015.134.
[34] J. Gao, E. C. Kara, S. Giri, and M. Berges, "A feasibility study of automated plug-load identification from high-frequency measurements," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2015, pp. 220–224, doi: 10.1109/GlobalSIP.2015.7418189.
[35] H. Liu, H. Wu, and C. Yu, "A hybrid model for appliance classification based on time series features," *Energy Buildings*, vol. 196, pp. 112–123, Aug. 2019, doi: 10.1016/j.enbuild.2019.05.028.
[36] S. Makonin and F. Popowich, "Nonintrusive load monitoring (NILM) performance evaluation," *Energy Efficiency*, vol. 8, no. 4, pp. 809–814, Jul. 2015, doi: 10.1007/s12053-014-9306-2.

KEXIN LI received the B.E. degree in computer science and technology from Anhui Normal University, China, in 2018. She is currently pursuing the professional master's degree with the Ocean University of China. Her current research interests include signal processing and acoustic data analysis.

BO YIN received the Ph.D. degree from the Department of Computer Science and Technology, Ocean University of China, Qingdao, China, in 2006. He is currently an Associate Professor with the Ocean University of China and a Visiting Scholar with Carnegie Mellon University and the University of Nantes, France. His research interests include acoustic systems, embedded system designing, and intelligent control technology.

ZEHUA DU received the M.S. degree from the Department of Information Science and Technology, Shandong University of Technology, Zibo, China, in 2016. He is currently pursuing the Ph.D. degree with the Ocean University of China. His research interests include signal processing, artificial intelligence, and other related research fields.

YUFEI SUN received the B.E. degree from the Department of Information Science and Technology, Qingdao University, Qingdao, China, in 2018. He is currently pursuing the M.E. degree with the Ocean University of China. His research interests include fatigue driving, machine learning, and other related research fields.