Ad Hoc Networks 123 (2021) 102643

Contents lists available at ScienceDirect

Ad Hoc Networks

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

A reliable deep learning-based algorithm design for IoT load identification in smart grid

Yanmei Jiang [a,d], Mingsheng Liu [a], Hao Peng [b,*], Md Zakirul Alam Bhuiyan [c]

[a] State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, 300130, China
[b] School of Cyber Science and Technology, Beihang University, Beijing, 100083, China
[c] Department of Computer and Information Sciences, Fordham University, NY, 10458, USA
[d] Department of Rail Transportation, Hebei Jiaotong Vocational and Technical College, Shijiazhuang, 050035, China

ARTICLE INFO

Keywords:
Deep convolution neural network
Feature discrimination
Load identification
Long short-term memory network
Non-intrusive
Time-frequency transformation

ABSTRACT

In IoT load monitoring system of the smart grid, the non-intrusive load monitoring and identification (NILMI) has become the research focus. However, the existing researches focus on the accuracy of load identification, neglecting the effectiveness of data sampled, the distinction of load abstract feature representation, and the reliability of load identification model. This paper proposes a novel algorithm framework of spatial–temporal convolution neural network for NILMI, namely DST-CNN, to realize the fine-grained load identification. In the DST-CNN framework, to ensure the accuracy and reliability of data usage, an signal enhancement method, AM-PCA, is used. To enhance the distinction of load abstract feature representation, an extraction mechanism of the spatial–temporal features is developed, which utilizes deep convolution networks and time-series recurrent neural networks (RNN). To improve the accuracy and reliability of load identification model, a hierarchical load classification mechanism is constructed, and the deep long short–term memory (LSTM) structure as the classifier. A considerable amount of the high-frequent current signals are sampled to validation the performance of the proposed method. The experimental results demonstrate the good generalization performance and superiority for NILMI in IoT load monitoring system.

1. Introduction

With the rapid development of Internet technology, IoT technology is widely used in the smart grid. In the IoT load monitoring system, NILMI is a critical part of the load-side management of the ubiquitous smart grid [1]. It could not only help enhance the operational efficiency of the power grid side, but can alleviate the energy pressure, and improve energy efficiency [2,3]. Especially, monitoring the current, voltage and the power of the device can effectively analyze the power consumption, the running situation of equipment, and the electricity usage of people [4–7]. However, most of the existing approaches for NILMI still need to be improved: the generalization of load identification model, the accuracy and reliability of load identification, and the distinction of load abstract feature representation. These can inhibit the performance or impact the reliability of the NILMI. In this article, our study mainly includes: improving the validity of load data sampled, the capability of load characteristic representation, and the accuracy of the fine-grained load identification without overlay scene. It is very important for improving the dependability of IoT load monitoring system in smart grid [8,9].

A substantial amount of methods for NILMI have be presented in the literature. These methods can be divided into two categories: the mathematical optimization and the data driving [10,11]. The former one is to solve the minimum of the combination results, namely the difference is between the total power of the equipment combination and the target loads value of a given set of equipment, and then in the combination, the equipment included is the identification result. The most of algorithms include the maximum likelihood function, the mean value and variance based on statistical characteristics, etc [12,13]. In the past few years, these methods have presented competitive performance for NILMI, however, there are still some shortcomings: relying on a great deal of prior knowledge; a huge amount of computations are required to determine the model parameters; the early models have some serious limitations, which can affect the reliability of IoT load monitoring system [14].

In the present stage, the data-driven methods include Decision Trees, k-Nearest Neighbor (kNN), Hidden Markov Models (HMM), and Random Forest (RF), etc, which have the distinct advantages over the previous methods of NILMI [15,16]. However, these methods require a

* Corresponding author.
E-mail address: penghao@buaa.edu.cn (H. Peng).

https://doi.org/10.1016/j.adhoc.2021.102643
Received 27 March 2021; Received in revised form 20 June 2021; Accepted 30 July 2021
Available online 8 August 2021

large number of offline data sets to train model; the parameter values need to be adjusted frequently and to identify the best model, which will increase the computational burden of the algorithm.

Recently, in view of the excellent feature representation ability, the deep learning has been quickly applied in NILMI. In [17], the author uses the convolution neural network (CNN) to extract the load features of the different appliances in NILMI. In [18], the Bayesian non-parametric learning-based approach and long short–term memory (LSTM) neural networks are used for load identification. There are a lot of the important information in the load features [19,20], so, many researches focus on improving the capability of load feature representation [10]. In [21], by extracting the voltage–current (V–I) trajectory of the household appliances as the load characterizes to realize the load monitoring. In [8], the current-to-image is extracted as the load feature from the residential appliance, and then the CNN is used as the load classifier to realize load identification. However, the shapes of V–I trajectory from a variety of appliances are similar, which significantly reduces their ability for load discernment. To improve the method, in [22], the author uses the short-time Fourier transform(STFT) and CNN to produce an image-like representation of load; however, the sampling error of load transient data is large, which requires the higher precision of experimental instrument; the hyper-parameters of model are set by engineering empirical operation, and it results in unstable performance.

A great deal of literature show that the deep learning methods are extraordinarily effective for improving the accuracy of load identification. In [23], the deep neural network (DNN) model structure is used to achieve the efficient monitoring and recognition of the running state of the appliance in NILMI. Considering the load characteristics are based on the time-series, some other literature utilize the LSTM model to realize the load identification. In [24], the LSTM-RNN is employed to enhance the accuracy of the single-load or multi-load appliance classification. However, such methods have their limitations: firstly, the generalization and global optimization of these methods are insufficient; secondly, in the real world, the optimal classifier for single-load identification is not for the multi-load; thirdly, increasing the number of layers of neural network to improve the performance of recognition model, which will lead to over-fitting of the model.

To address aforementioned problems, in this article, a novel algorithm framework based on deep learning, namely DST-CNN, is proposed. Our research goal is to present a high-reliability algorithm framework to realize the high-level load feature representation and the fine-grained accuracy of load identification for MILMI. In the framework, an effective method of the signal enhancement, namely the principal component analysis based on adjacency matrix (AM-PCA), is used to reduce the error of data collected and ensure the effectiveness of data. To improve the capability of load representation, the time–frequency transformation technology based on the short-time Fourier transform(STFT) is used. The DNN and time-series recursive neural networks (RNN) are used to form a mechanism of spatial–temporal feature extraction, and to obtain the high-quality load feature mapping. The long short-term memory networks (LSTMs) method as the load classifier is use to realize the fine-grained hierarchy of load identification. The real-time experimental data sets are used to verify the performance of the proposed model, and the results present the significant reliability of the model for IoT load monitoring system. Fig. 1 shows the framework of DST-CNN in the non-intrusive load monitoring system of IoT.

From Fig. 1, the smart grid consists of the power side, grid side and load side. The IoT load monitoring system is the key component of the load side. The non-intrusive load monitoring system of IoT is used to identify the load type and analyze the power consumption of the equipment in real-time. The perception layer as the load data provider can collect the current, voltage and active power signals of various appliances, which can be collected by the smart meters installed at the entrance to buildings in IoT. The DST-CNN is the heart of MILMI algorithm.

The main contributions of this paper are as follows:

(1) The paper uses an effective method of signal enhancement based on AM-PCA to achieve de-noising of the raw data, and ensure the validity and reliable of data sets.
(2) Utilizing STFT technology to realize time–frequency domain transformation of the steady-state current; using the DNN model to improve the ability of load spatial representation.
(3) Using a reconstructs RNN as the feature extractor of the load spatial–temporal, which can provide the high-quality load features mapping; the deep LSTM-RNN based on the time-series as the load classifier, and the method can improve the performance of load identification with the fine particle size.

The other chapters of this paper are arranged as follows. Section 2 reviews the related work. Section 3 introduces the algorithm design details of DST-CNN model, and the evaluative criteria of the proposed model. Section 4 presents and discusses the experimental setting of the model, the experimental results, and the implementation of the model. Section 5 concludes this article.

2. Related works

This chapter briefly reviews the learning of load event detection, load feature extraction based on the frame frequency-domain image, and the spatial-temporal representation based on deep learning.

2.1. Load event detection

In NILMI system, there are two types of load characteristic events: the transient and the steady state [25]. The transient state last for a short time and easily subject to large signal fluctuations, which affects the accuracy of load event acquisition [26,27]. Considering the types and the loads time-domain waveform of household appliances are relatively regular, this paper obtains the high-frequency steady-state load signal for the event detection.

As relatively straightforward parameters, the current and voltage are monitored to describe the changes in running state of household appliance, and they can be obtained by the simple technique [28]. In general, the zero crossing detection (ZCD) is used to determine the period of the data sampled [29,30]. When the running state of load is stabilized after load start-up or state transition, the current is defined as the steady-state current. By a lot of experimental measurements, the stable-state of the most electrical appliances is determined after appliances starting up or state transition for 2s. So the time-domain waveform of the steady-state current is collected after appliance has been working for 2s in the NILMI system [31]. One-cycle steady-state current waveform is measured at a position where the steady-state voltage waveform crosses zero and rises. (j) and $(j+1)$ represent the continuous sampling point-in-time [30]. Fig. 2 shows one-cycle steady-state current and voltage waveform from micro-wave oven, where $u(j)$ and $u(j+1)$ represent the voltage signal of the (j) and $(j+1)$ moment sampling point respectively, and the voltage waveform crosses zero and rises.

2.2. Load feature representation based on frame frequency-domain image

The feature extraction of load is the important component of NILMI system. The ability of the abstract feature representation directly affects the reliability of load identification. In general, the fast Fourier transform (FFT), wavelet analysis, and the shape characteristics of the voltage–current (V–I) trajectory are used for identifying the single device with a switching mode power supply, but the anti-interference ability and generality are poor [32]. The wavelet analysis can handle overlapping events to a certain degree, however, it is not suitable for long-term operation of load equipment [33]. The voltage–current (V–I) trajectories are widely used for abstracting the load feature to enhance load discrimination [6,34]. However, these methods hard to distinguish

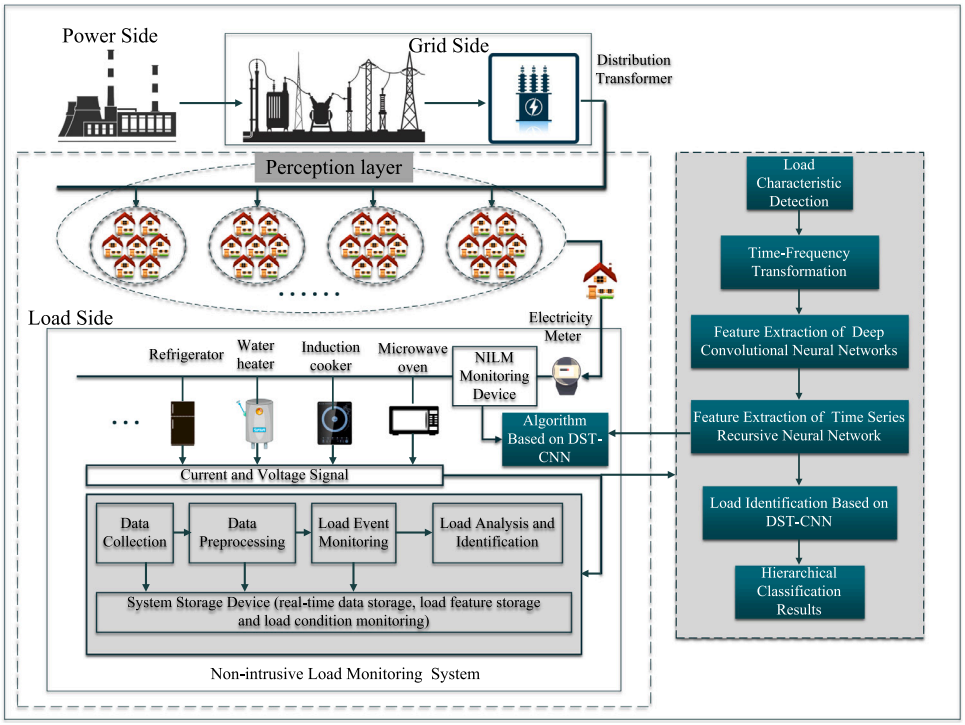

Fig. 1. The framework of DST-CNN for non-intrusive load monitoring and identification.

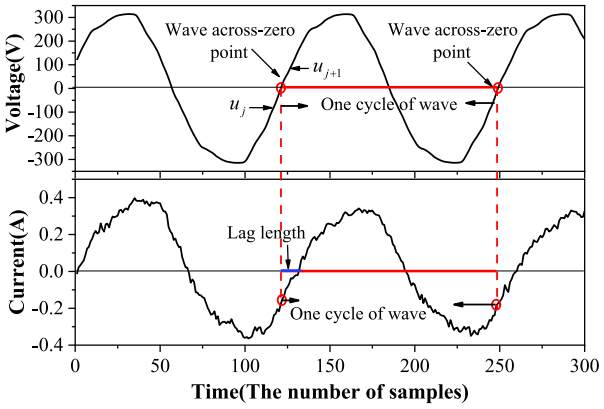

Fig. 2. The one-cycle of current and voltage waveform from micro-wave oven.

between the small-power devices with the similar V–I shape. Enlightened by the remarkable effect of the frame division image in the image recognition. In this paper, we select the frame frequency-domain image as load characterization [35]. The short-time Fourier transform (STFT) is employed to realize the time–frequency conversion of the steady-state current waveform, and then to split the frequency image by frame. The segmented frequency-images are as the pre-processed data for the spatial–temporal feature extraction of load.

2.3. Deep spatial–temporal representation of load

Load feature learning has become the research focus in NILMI. The accuracy of load feature representation affects the performance of load monitoring and identification. Using deep learning technology for load feature extraction aims at improving the quality of discriminative feature. Artificial Neural Networks (ANNs) [36], convolutional neural networks (CNNs) [37], Hidden Markov Model (HMM) [38], and the support vector machines (SVM) [39], etc, have been widely used in NILMI. These methods can simply abstract the essential characteristics of the load, however, ignoring the load has the characteristic of the time series. Long short–term memory networks (LSTMs) has good performance in dealing with the long time dependencies of data, but it has limited ability of the load feature extraction [40]. Especially, in the multi-load feature recognition, the incomplete feature extraction results in the identification inaccuracy. Meanwhile, the over-fitting and non-convergence of loss function for identification are to be considered.

The traditional approach based on the time series, such as Autoregressive Integrated Moving Average (ARIMA) [41,42]. It can well represent the connection between the load sampling data points at a certain moment, however, in the process of modeling, this method has large fluctuations, strong uncertainty and the poor stability of the model. The spatial–temporal representation learning is attractive method to tackle the feature extraction problem based on time series. In [43], utilizing the spatial–temporal representation to realize the effective and full captures for the features time correlations, meanwhile the method can present the good estimated performance for the events forecast. In [44], a simple spatial–temporal pattern for energy decomposition is used. However, considering the complexity of time series data sampled in multivariate load, the performance of the model still need to be improved. As the network scale becomes larger, the problem of model optimization can be difficult to solve. In this paper, the deep spatial–temporal representation is employed to abstract the high-quality load features, and to realize the fine-grained load identification.

3. Deep spatial–temporal convolution neural network for load identification

In this section, the algorithm framework of DST-CNN model is proposed to address the following research problems: the effectiveness of sampled data, the incomplete load characteristic representation, and

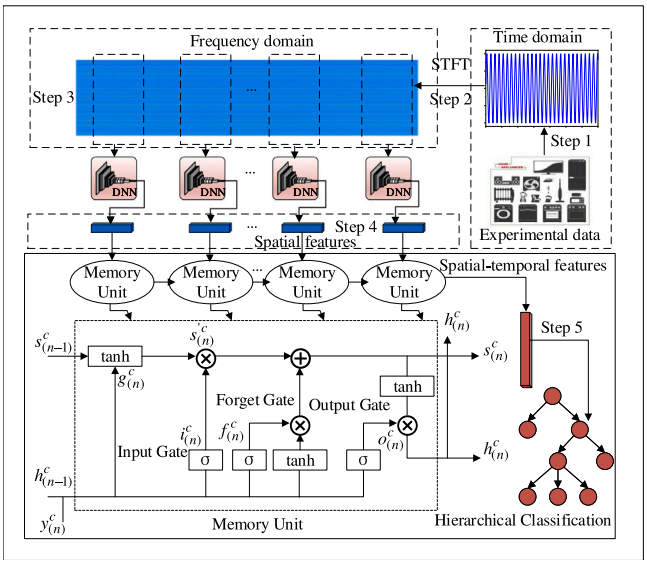

Fig. 3. Block diagram of the developed DST-CNN model.

the accuracy and reliability of load identification. The overview of model is shown in Fig. 3. We mainly introduce the load event detection and data acquisition, load feature discrimination based on the time–frequency transformation, spatial feature extraction mechanism based on DNN, and spatial–temporal feature extraction mechanism based on LSTM-RNN.

3.1. The key design steps of the DST-CNN model

In our work, the design of the DST-CNN model includes 5 steps. In step 1, the experimental data of the steady-state current and voltage signals are sampled from the smart electricity meters installed in household. In step 2, the STFT is used to realize the time–frequency transformation of current collected; in this process, the time domain waveform of the steady-state current to be transformed into the frequency–domain images as the abstract characteristics of load. In step 3, the concept of frame is introduced to the frequency-domain images divided; an integral frequency-domain image is divided into many single frame images. In step 4, each single frame image as the input vector of the DNN is inputted into DNN model for convolution and pooling operations, finally to generate the spatial features sequence of load features mapping. In step 5, each spatial feature is inputted into the temporal recursive neural network respectively; the spatial–temporal features mapping are generated for the fine-grained type identification of loads, and to achieve the hierarchical classification of load. The detailed design of each part of the model and the meaning of parameters will be introduced later in this section.

3.2. Load event detection and data acquisition

In this paper, compared to the low-frequency sampled data, the high-frequency sampled data can provide the sufficient load features information, which is very important to improve the capability of features representation learning in NILMI. [45]. The one-cycle of load current signal is determined by the ZCD of voltage, which as a simple and effective way has been used in some literature. [46,47]. This process is shown in Algorithm 1. When the steady-state voltage waveform at least two crosses zero and rises, and the current signal corresponding to the voltage signal between the two zero crossings are collected [30].

Algorithm 1: Load event detection

Input: U, I, Samplying points size M;
Zero crossings: voltage C_{zu}, current C_{zi};
Sampling size of one complete cycle CT;
Output: $i_m, u_m, (m \in M)$;
Initialized data: i_0, u_0
//Detect the voltage zero-crossings;
Obtain the current zero-crossings;
for $m = 2$ to $len(C_{zu}) - 2$ do
$CT_m = C_{zu}[(m+2)] - C_{zu}[(m)]$;
$u_m = U[C_{zu}(m) : C_{zu}(m+2)]$;
$i_m = I[C_{zu}(m) : C_{zu}(m+2)]$;
if $CT_m = CT$ and $C_{zi} \geq 2$ and $len(i_m) = CT$ **then**
break;
endif
endfor

In the paper, the four common kinds of household appliances were selected such as refrigerator (Refrig), water heater (WH), induction cooker (Inco), and microwave oven (Micro) for load recognition.

3.3. Data pre-processing with AM-PCA

In the real work, the inevitable instrumentation error does exist, so a signal de-noising method, the principal component analysis based on adjacent matrix (AM-PCA) is employed. The method utilizes a neighbor current waveform signal to amend the error sampling point for reducing error in the sampling waveform signal [48]. The similar methods have been used in other fields [49]. Fig. 4 shows the process flow diagram of AM-PCA.

The mathematical expression of adjacent matrix is formulated as

$$D = \begin{bmatrix} d^1_{(i-1)} & d^1_{(i)} & d^1_{(i+1)} \\ d^2_{(i-1)} & d^2_{(i)} & d^2_{(i+1)} \\ \vdots & \vdots & \vdots \\ d^N_{(i-1)} & d^N_{(i)} & d^N_{(i+1)} \end{bmatrix} \tag{1}$$

where, the $(i-1)$, (i), and $(i+1)$ are a continuous time series. The $d^n_i (n \in N)$ represents the nth sample of ith sampling period, and the sampling point d^n_{i-1} and d^n_{i+1} adjacent to it. The $d(i-1)$, $d(i)$, and $d(i+1)$ are a sequence of sample points at a continuous time. The matrix D represents a current signal matrix. The number of sample in each sampling cycle is N.

In Fig. 4, the process of AM-PCA is that three adjacent current signals sequences, $d(i-1)$, $d(i)$, and $d(i+1)$ are selected, and then by extracting the principal component of $d(i-1)$ and $d(i+1)$ to enhance the signals sequences $(d^1_i, d^2_i, \ldots, d^n_i)$ at ith moment. The singular value decomposition (SVD) is used to optimize the performance of PCA algorithm. $d_n = d_n - \frac{1}{m}\sum_{j=1}^{m} d_j$ represents the decentralization of the sampled sequences. Calculating the principal component of the n' dimension of the sample d^n_i is equivalent to calculate the eigenvector matrix W of the covariance matrix XX^T of the sample set, which is correspond to the first n' eigenvalues. Finally, each sample d^n_{i-1}, d^n_i, and d^n_{i+1} will be transformed as $z^n_{i-1} = W^T d^n_{i-1}$, $z^n_i = W^T d^n_i$, and $z^n_{i+1} = W^T d^n_{i+1}$; the updated sample set $D' = (z^1, z^2, \ldots, z^n)$ are obtained. The method can be used for detecting the error sampling points.

3.4. Load feature discrimination with STFT

In general, the features of load signals from multi-sensor system are high dimensional, which contains some redundant and noisy. The signal dimension reduction can improve data efficiency. To avoid the loss of the majority of the current characteristics in dimension reduction and ensure the reliability of the load characteristics extracted, the short-time Fourier transform (STFT) is employed. As a practical technology for speech signal processing, STFT uses a distribution class of time and

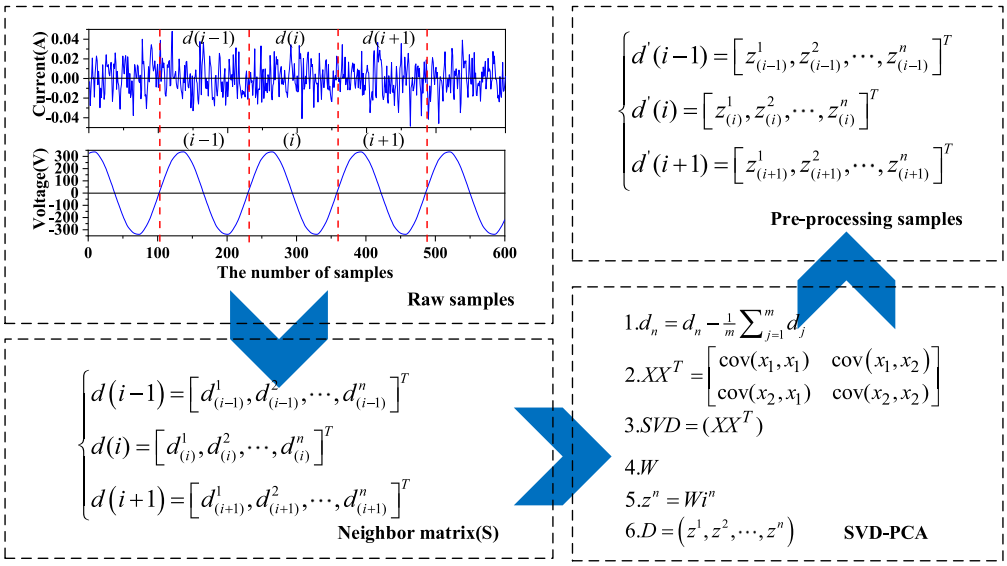

Fig. 4. The process flow diagram of AM-PCA.

frequency to specify the complex amplitude of any signal varies with time and frequency [50]. The STFT can be defined as below:

$$STFT_z(t,f) = \int_{-\infty}^{\infty} [z(u)g^*(u-t)]e^{-j2\pi fu}\mathrm{d}u \tag{2}$$

where $z(u)$ represents the source signal, $g^*(t)$ represents the window function, t represents time. The function is described as the STFT of signal $z(u)$ at t moment, which is the Fourier transform of the signal $z(u)$ multiplied by an analysis window $g^*(u-t)$ that is the center of t moment. Herein, $z(u)$ represents the time-domain waveform of the steady-state current. $STFT_z(t,f)$ represents the frequency spectrum at t moment.

Fig. 5 shows the time-domain waveform of the steady-state current signal from 4 different types of appliances, and Fig. 6 shows the converted frequency-domain images of the steady-state current signal. From Fig. 6, the frequency-domain images of the different appliances can represent the different load characteristics of the appliances.

3.5. Spatial features extraction of DST-CNN model

As the key component of DST-CNN model, DNN is employed to realize the spatial features extraction of the household appliances load. Fig. 7 demonstrates the schematic diagram of the load spatial features extracted. Briefly, the algorithm structure of DNN mainly includes two mainly parts: (1) features extractin from the frame image; (2) the construction of DNN structure for the spatial features extraction.

Step 1: feature extraction based on DNN. After the frequency-domain images of load have been obtained, the complete frequency-domain image of load is divided into some smaller frame images, and then form a sequence of single frame images. In [51], the result of the experiment confirms that a short-time property of frame feature is better than a long-time distributional frame pattern. The process is as follows: we sample a frequency-domain image of load, the sequence X with M frames; by this way, the sequence frame is defined as $R = \sum_{r=1}^{M} R_i$, where, R_i is a frame, the value of i is from 1 to M. The each frame R_i as a feature vector is inputted to the first convolution layer of DNN, and then through the convolutional kernel layer to form the features mapping. Each image of the sequence of frame images has the temporal characteristics, and each frame image of the sequence frame images as a highly discriminative feature.

Step 2: construction of DNN structure for spatial features extraction. In this paper, to form the spatial feature representations of load, a deep convolution neural network (DNN) is proposed. The structure of DNN includes the convolution network layer, the max-pooling layer, and the full-connection network layer. The number of parameters in each part is set to nine, six, and four respectively. Between the convolution layer and maxi-pooling layer, we select rectified linear unit (ReLU) as the activation function. ReLU can boost the nonlinearity of function to avoid gradient descent. The expression of ReLU is as follows:

$$\mathrm{ReLU}(x_i) = \max(0, x_i) \tag{3}$$

The feature map of the new convolution layers could be calculated as

$$\begin{cases} C_i^n = \sigma(\sum_{s\in M_{i(n-1)}} \sum_{(p,q)\in K^{(n)}} w_{is(p,q)}^s x_s^{(n-1)}(c+p, r+q) + b_i^{(n)}) \\ K^n = [(p,q) \in N^2 | 0 < p < k_w, 0 < q < k_h] \end{cases} \tag{4}$$

where, $K^{(n)}$ represents convolution kernel, n indicates the number of the CNN layer, the size of convolution kernel is $K_w \times K_h$, the height is K_h, and the weight is K_w; $b_i^{(n)}$ is the bias of the ith characteristic mapping in convolution network layers, $w_{is(n-1)}^s$ is the weight of the neuron of the ith feature mapping in convolution layers, s is the previous state, the parameters c and r are the column and row of the input image respectively. The function σ represents the activation function ReLU. After the convolution layers, the features of image are inputted in the max-pooling, and to reduce overfitting. This detailed process can be formulated as:

$$Y_i^n = down(C_i^n) \tag{5}$$

The $down(\bullet)$ represents the down-sampling function, C_i^n indicates the features mapping. The max-pooling layer can reconstitute the new features mapping with the max value of a filter $n \times n$. In this process, if the inputed frame image is $r_i(x,y)$, after it through the first convolution kernel $k_i(p,q)$ with the size $a \times b$, the corresponding feature mapping is $m_i(x,y)$, which can achieved the calculation by Eq. (4). After the convolution layer, the feature mapping $m_i(x,y)$ will be inputted into the max-pooling, and the size of the feature mapping transfers into $m_i'(x,y)$ by Eq. (5), and then the features are extracted and fed into a fully-connected layer. Finally, the spatial features mapping are formed, $sm_i'(x,y)$, which is one-dimensional array, which will as the input layers are fed into the LSTM-RNN model.

3.6. Spatial–temporal feature extraction of DST-CNN model

In this section, we present a LSTM-RNN model to form the spatial–temporal feature mapping of the electricity load for the fine-grained load identification. The structure of LSTM-RNN model is shown in

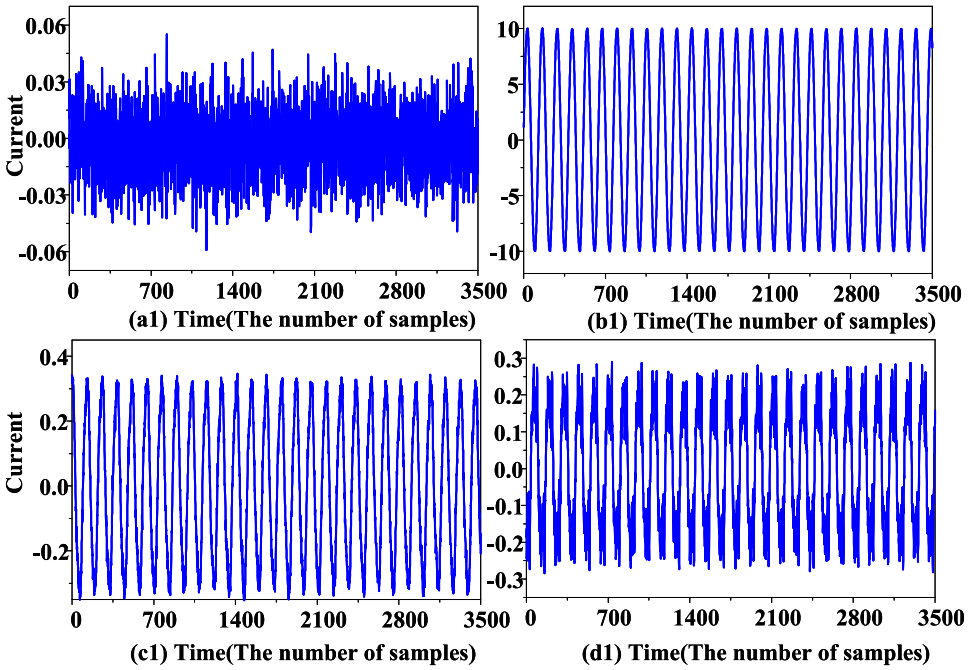

Fig. 5. The time-domain steady-current waveform of 4 different appliances.(a1) Refrigerator. (b1) Water Heater. (c1) Micro-wave Oven.(d1) Induction Cooker.

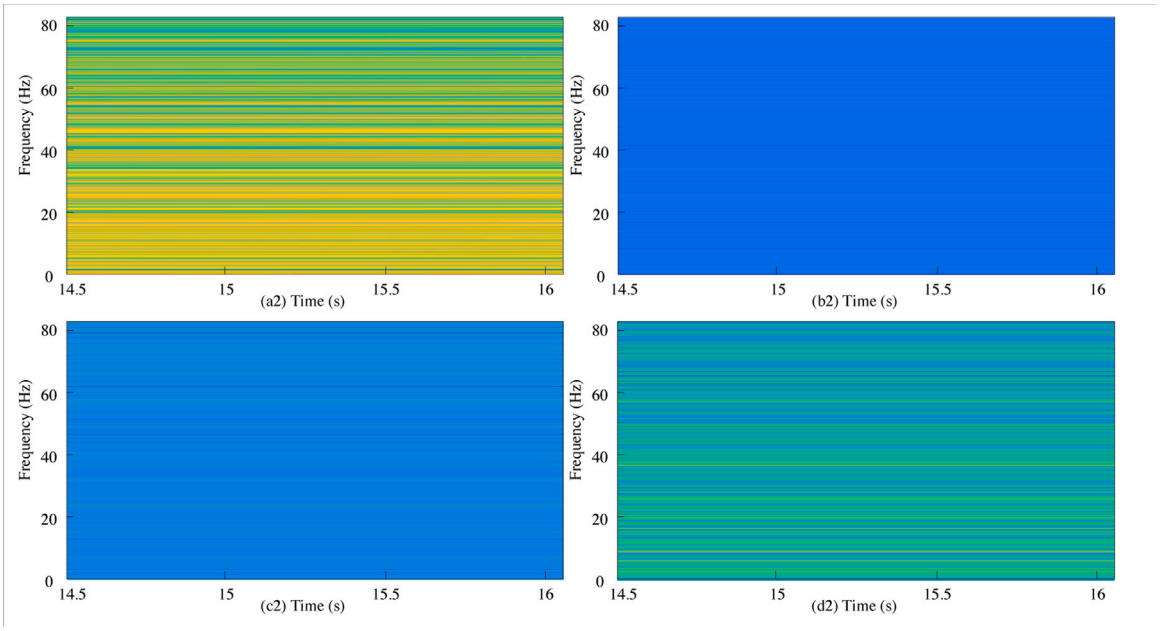

Fig. 6. The frequency-domain images of 4 different appliances. (a2) Refrigerator. (b2) Water Heater. (c2) Micro-wave Oven. (d2) Induction Cooker.

Fig. 8. LSTM-RNN is used to solve the long dependency problem of network. In the paper, since the load has the characteristics of time series, LSTMs is used for the spatial–temporal feature extraction of the load. The detailed structure of the LSTMs is presented in Fig. 9.

The core of LSTMs model includes an input gate, a forget gate, and an output gate. The superscript c of each vector represents the neuron, the subscript n represents time, $y_{(n)}$ represents the n moment input layer, and $h_{(n-1)}$ represents the $(n-1)$ moment hidden layer [52]. The vector $a^c_{(n)}$ is in the n moment, the current state of LSTMs, and the center of each neuron has been linearly activated. Forget gate $f^c_{(n)}$ is the first step in the memory unit of LSTMs, which selectively delete certain information of state. In the moment, $f^c_{(n)}$ is used to controlled information of the input layer, and $h^c_{(n-1)}$ is the previous moment of the hidden layer, and they determine the input of $f^c_{(n)}$. The expression of $f^c_{(n)}$ can be formulated as follows:

$$f^c_{(n)} = \sigma(W_f \cdot [h^c_{(n-1)}, y^c_{(n)}] + b_f) \tag{6}$$

The input gate and input node. The $i^c_{(n)}$ represents the input gate of the LSTMs, and it is a sigmoid function layer to decide which value will be updated. The $g^c_{(n)}$ represent the input node, and it is a tanh function layer to create the candidate state vector $a'^c_{(n)}$ states. The $i^c_{(n)}$ is related to $h^c_{(n-1)}$ and $y^c_{(n)}$. The calculation formula of $i^c_{(n)}$ is defined as

$$i^c_{(n)} = \sigma(W_i \cdot [h^c_{(n-1)}, y^c_{(n)}] + b_i) \tag{7}$$

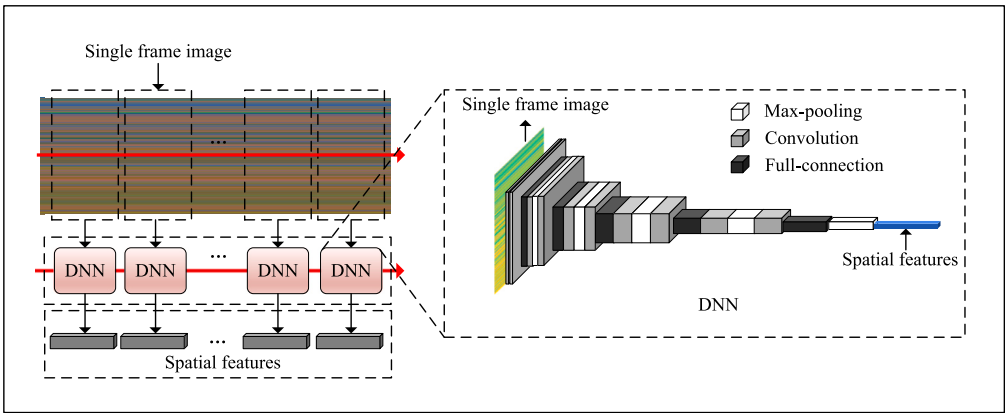

Fig. 7. The schematic diagram of the load spatial features extracted.

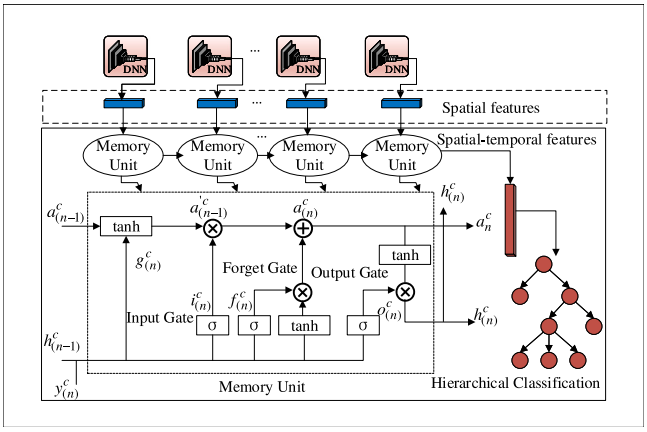

Fig. 8. The structure of LSTM-RNN model.

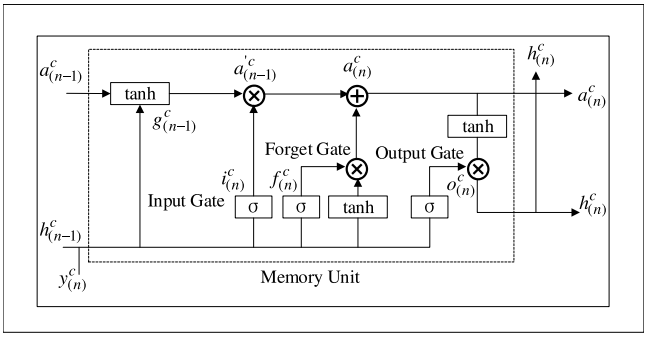

Fig. 9. The memory unit structure of LSTM.

The candidate state vectors $a'^{c}_{(n)}$ is defined as follows:

$$a'^{c}_{(n)} = \tanh(W_a \cdot [h^c_{(n-1)}, y^c_{(n)}] + b_a) \tag{8}$$

The input node $g^c_{(n)}$ is defined as

$$g^c_{(n)} = \tanh(W_g \cdot [h^c_{(n-1)}, y^c_{(n)}] + b_g) \tag{9}$$

The update of the current status $a^c_{(n)}$ is represented as

$$a^c_{(n)} = g^c_{(n)} * i^c_{(n)} + a^c_{(n-1)} * f^c_{(n)} \tag{10}$$

The $o^c_{(n)}$ represents an output gate of the model, and it is used to determine what information to be outputted. The internal state a^c utilizes a tanh layer to obtain the value of the state $a^c_{(n)}$, and the value is between −1 and 1, then multiply by the output of sigmoid function layer to get the rest of the state. The detail description of $o^c_{(n)}$ and $h^c_{(n)}$ are given as:

$$o^c_{(n)} = \sigma(W_o \cdot [h^c_{(n-1)}, y^c_{(n)}] + b_o) \tag{11}$$

$$h^c_{(n)} = \tanh(a^c_{(n)}) * o^c_{(n)} \tag{12}$$

where w and b respectively represent the layer weight value and the offset value.

The output layer of LSTMs uses the softmax to realize the hierarchical classification of the load, and the softmax can easily implement the multiple types of nonlinear classification. The detailed process is described as follows:

$$S(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{M} e^{z_j}} \tag{13}$$

where, z is the output of the full-connected layer, and $\sum_{j=1}^{M} e^{z_j}$ represents the normalizes the output probabilities. The softmax turns the output of LSTM into a probability distribution, and the cross-entropy is used to calculate the distance between the predicted probability distribution and the true data probability distribution. Loss of the cross-entropy is defined as

$$L_{class} = -\sum_{i=1}^{N} y_i log S_i \tag{14}$$

In the paper, the DST-CNN model is employed to realize the fine-grained load identification in NILM system. The whole process of DST-CNN is shown in algorithm 2.

Algorithm 2: The whole process of DST-CNN

Input: Initial data$[I_0, I_1, \ldots, I_{(t-1)}]$;
Pre-processing data: $[I'_0, I'_1, \ldots, I'_{(t-1)}]$;
Feature data: $[F'_0, F'_1, \ldots, F'_{(t-1)}]$;
Number of feature:N;Number of epochs: NE; Batches: NB.
Output: The DST-CNN model M
// Train the model
for $ne = 1, 2, \ldots, NE$ **do**
for $nb = 1, 2, \ldots, NB$ **do**
$z_i \leftarrow$ Eqs. (4) and (5);
$f_{(n)} \leftarrow$ Eq. (6);
$i_{(n)} \leftarrow$ Eq. (7);
$g_{(n)} \leftarrow$ Eq. (9);
$o_{(n)} \leftarrow$ Eq. (11);
$h_{(n)} \leftarrow$ Eq. (12);
$s(z)_i \leftarrow$ Eq. (13);
$L_{class} \leftarrow$ Eq. (14);
end
end

3.7. Evaluation criterion of DST-CNN model

In our experiment, some practical evaluation criteria are employed for the performance of the proposed model, namely the model accuracy, Matthews correlation coefficient (MCC), Zero-loss score (ZL), and Macro $-F_1$. The average value of all the precision is Macro $-F_1$ score, the formula is describe as

$$\text{Macro}-F_1 = \frac{i}{M}\sum_{m=1}^{M} F_{1(m)} \tag{15}$$

where, the number of appliances is M, $F_{1(m)}$ is a weighted average of model precision and recall rates of mth appliance, the maximum is 1, and the minimum is 0. The $F_{1(m)}$ is defined as

$$F_{1(m)} = 2 \cdot \frac{p \cdot r}{p + r} \tag{16}$$

where p and r represent the precision and the recall respectively.

Zero-loss score (ZL) represents the 0_1 loss of classification, and the calculating formula is deduced as

$$L_{0-1}(x_k, \hat{x_k}) = 1(\hat{x_k} \neq x_k) \tag{17}$$

where assuming $\hat{x_k}$ is the predictions of kth appliance, x_k is the true value of kth appliance.

MCC is called a balanced metric, and it can be used for classes with different sizes; MCC is a relational value between −1 and 1. If the correlation coefficient is -1, and illustrates a complete inconsistency between prediction and observation; if the value is 1, and illustrates a perfect prediction, 0 shows an random prediction. MCC can be defined as follows:

$$MCC = \frac{tp_m \times tn_m - fp_m \times fn_m}{\sqrt{(tp_m + fp_m)(tp_m + fn_m)(tn_m + fp_m)(tn_m + fn_m)}} \tag{18}$$

In general, tp_m, tn_m, fp_m, and fn_m represent the true positive, true negative, false positive, and false negative of mth class respectively.

4. Experimental setups

This chapter presents that the real measured datasets are utilized to train, test, and validate the proposed method. For better illustration about the performance advantage of the proposed method, three other typical methods of load recognition are introduced. The simulation environment configuration includes Python 3.5, the PyTorch 1.0, and the NVIDIA CUDA 9.0. Python 3.5 is used to program for the proposed algorithm implementation.

4.1. Experiment datasets

In this paper, the experimental data acquisition are from 50 buildings and the data collecting period is between 2018 and 2020. The time spent on using household appliances of one day is 24 h (sampling frequency $f_s = 130$). The data mainly includes four types of the most representative household appliances, namely refrigerator (Refrig), water heater (WH), microwave oven (Micro), and induction cooker (Inco). These data contain the steady-state voltage, current and active power of the high-frequency load, which can provide the sufficient experimental data sets for NILMI.

There are two key reasons to explain why we select the above four kinds of electrical appliances to experiment. Firstly, the above selected experimental appliances are all the representative household appliances, and they are indispensable in modern life. Secondly, the changes in running state of those appliances are from simple to complex forms, and the variable running state samples are good for training models under different conditions. Thirdly, considering the limited space of the paper, the experiment only uses some common load combinations.

We use the relatively small data set No.1 building, No.2 building and No.3 building for experiment. There are 5×10^9 data samples collected over the course of a year. These data samples consist of two parts: the status data of the four kinds of electrical appliances working alone; the status data of the simultaneous running of two or more appliances. The sampling numbers of each type of appliances are 10^7 for training and testing model, which can provide the sufficient single and multiple operating waveforms data.

To ensure the effectiveness of the acquired dataset, this experiment takes the average of the three measurements as the final current value. Since existing the fluctuations of measuring signal, noise, and the abnormal spikes in sampling data, the AM-PCA is employed to complete the data pre-processing, which has been introduced in the previous section. In the paper, the 80% datasets are selected to train the model and the rest 20% to test the model.

4.2. The proposed model training and testing

For better illustration about the advantageous performance of the proposed method, the other three typical load recognition methods, namely KNN, CNN, and LSTM, are introduced to present the robustness of our method in the paper. The parameters of DST-CNN model are set as the number of the hidden layer blocks is 5, each block consists of 128 filters, and the filter size is 3. The pooling size and the pooling stride are all set to 2. ReLU is selected as the activation function, and the value of the dropout probability is set to 0.5. The size of mini-batch is set to 512, the Adam optimizer is generally preferred and learning rate is set to 0.002. The number of epoch is set to 200 with an early stopping criterion to prevent over-fitting.

In the experiment, we use the datasets from three buildings mentioned above to discuss the experimental results of our proposed method. We performed 5-fold validation on each of the datasets. The accuracy of training process are reported in Fig. 10, where the best results are shown. From Fig. 10, among the training process of above 4 models, the accuracy of DST-CNN method is the highest, and the mean accuracy is 91.64%, which is 9.40% higher than the second-best ranked CNN. The CNN method has advantages for the image recognition, but for the recognition of load feature image based on temporal characteristics, the accuracy is relatively low. The mean accuracy of KNN and LSTM is 70.59% and 68.49%, respectively. The LSTM is the lowest in four methods.

To further to analyze the performance of the proposed model, we used three testing dataset from three buildings to test the four model respectively, and we performed 3-fold validation on each of the datasets. By calculating the average value of the three verification results as the final results, the testing results are presented and discussed in Fig. 11. The results show that the testing accuracy of DST-CNN method is better than the other 3 models on the different testing datasets.

4.3. Evaluation criteria for the proposed model

To evaluate the performance of the proposed model, the single-label and the multi-label load identification are considered for our experiment. The Macro $-F_1$ scores, ZL scores, and MCC are employed to evaluate the four types of load identification methods in the single-label and the multi-label experiment. The experimental datasets are from the three different buildings, each dataset includes the single and the combined appliances, which are Refrig, WH, Micro, Inco, Re+WH, Re+Micro, Re+Inco, Re+WH+Inco, Re+WH+Micro, Re+Micro+Inco, Re+WH +Micro+Inco.

In Figs. 12 and 13, we use the three datasets from the different buildings to verify the Macro$-F_1$ scores of the four method. It is obvious that the DST-CNN method is better than the other three methods under all labeling conditions. In Fig. 12, it mainly demonstrates the Macro$-F_1$ scores of the single load identification test, and the Macro $-F_1$ scores of DST-CNN method are the best. The CNN method ranks the second, and better than the KNN and LSTM methods. In view of the time-series characteristics of the load, the results clearly illustrate that the KNN and LSTM methods are inferior to DST-CNN and CNN methods. Fig. 13

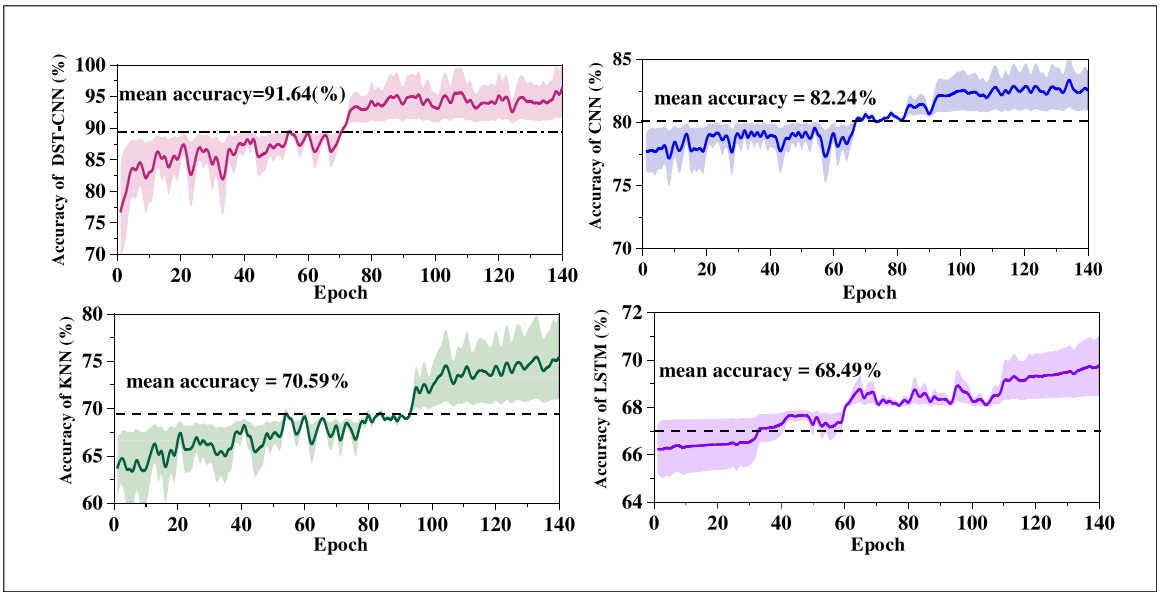

Fig. 10. The accuracy of training process.

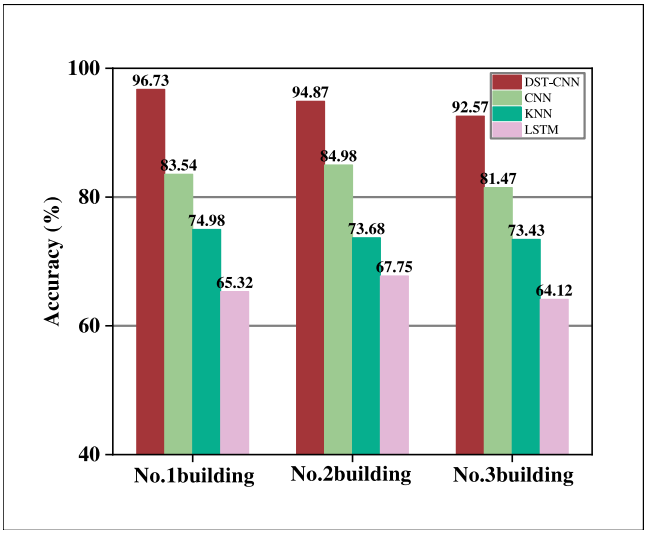

Fig. 11. The accuracy of testing process.

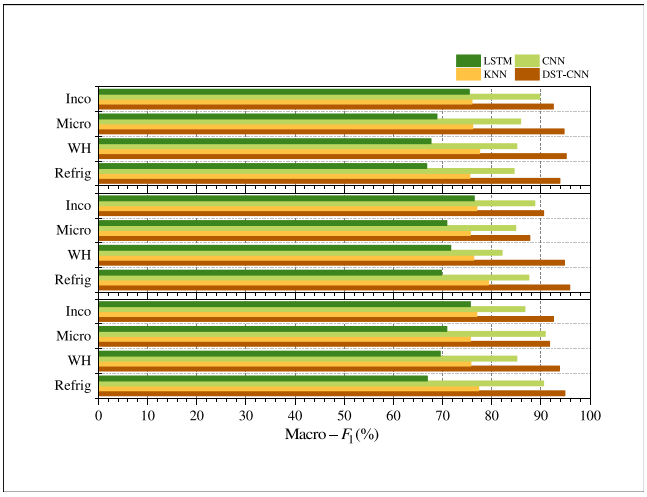

Fig. 12. Macro – F_1 scores of the single load identification test.

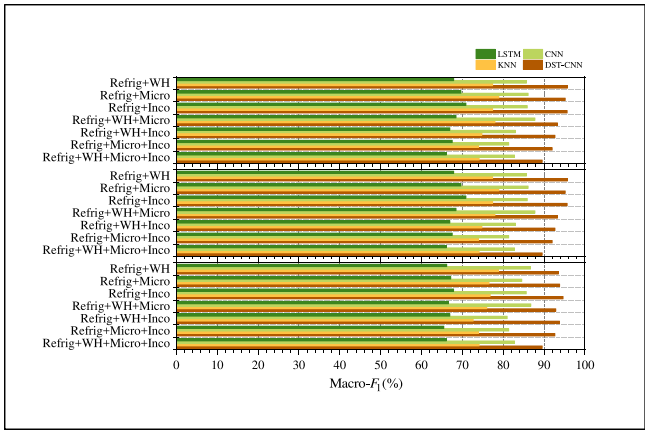

Fig. 13. Macro – F_1 scores of the multi-load identification test.

presents the Macro – F_1 scores of the multi-load identification test, and the Macro – F_1 scores of DST-CNN method is much better than the other three methods. In terms of the time series characteristic of load, the CNN method has no advantage in the multi-load recognizing. These results obviously indicate that the DST-CNN method is effective for the load signal identification of NILMI under all labeling conditions.

The ZL scores and the MCC as the performance indicators are also used to evaluate the proposed model with the four methods and the two different datasets from two buildings (No.1 buildings and No.2 buildings). In Tables 1, 2, and 3, the average scores and the standard deviations of experimental are clearly demonstrated, and the bold scores represent the best results. Table 1 presents the results of four methods on the single load datasets, which are from the No. 1 building and No. 2 building respectively. The results show that the ZL scores of DST-CNN method are less than the other three methods and the MCC scores are higher than the other three methods, indicating that the identification performance of DST-CNN method in the single load is better. However, the MCC value of CNN method is the highest for refrigerator identification, illuminating that CNN is an effective method for the single devices identification, and the case rarely occurs in the combinatorial appliances identification.

The ZL scores and the MCC scores of the multi-labels experiment of the appliance identification can more deeply assess the proposed

Table 1
The key evaluation indexes of single appliances.

Index	Method	No. 1 building				No. 2 building			
		Refrig	WH	Micro	Inco	Refrig	WH	Micro	Inco
ZL	DST-CNN	10.47 ±0.13	**1.18 ±0.021**	**6.54 ±0.161**	**6.23 ±0.173**	9.11 ±0.112	**5.32 ±0.124**	**5.30 ±0.091**	**5.97 ±0.013**
	KNN	27.01 ±0.062	24.45 ±0.038	18.99 ±0.012	19.01 ±0.043	21.13 ±0.052	25.14 ±0.089	19.13 ±0.045	19.18 ±0.012
	CNN	**6.30 ±0.036**	15.54 ±0.012	13.47 ±0.031	14.50 ±0.057	**6.27 ±0.014**	9.46 ±0.011	10.43 ±0.023	11.49 ±0.009
	LSTM	27.88 ±0.045	26.73 ±0.034	28.05 ±0.052	26.52 ±0.018	27.49 ±0.023	28.17 ±0.019	26.41 ±0.015	25.72 ±0.016
MCC	DST-CNN	0.88 ±0.035	**0.99 ±0.014**	**0.92 ±0.022**	**0.93 ±0.011**	0.90 ±0.014	**0.95 ±0.014**	**0.95 ±0.021**	**0.94 ±0.009**
	KNN	0.70 ±0.025	0.74 ±0.021	0.79 ±0.031	0.78 ±0.014	0.77 ±0.017	0.73 ±0.016	0.78 ±0.025	0.78 ±0.021
	CNN	**0.93 ±0.010**	0.82 ±0.009	0.85 ±0.021	0.84 ±0.045	**0.93 ±0.038**	0.89 ±0.051	0.88 ±0.047	0.87 ±0.031
	LSTM	0.69 ±0.023	0.70 ±0.031	0.68 ±0.015	0.70 ±0.014	0.69 ±0.053	0.66 ±0.063	0.71 ±0.062	0.72 ±0.013

Table 2
The key evaluation indexes of combined appliances for No. 1 building.

Index	Method	Refrig +WH	Refrig + Micro	Refrig + Inco	Refrig + WH + Micro	Refrig + WH + Inco	Refrig +Micro +Inco	Refrig + WH + Micro + Inco
ZL	DST-CNN	**1.17 ±0.023**	**6.25 ±0.019**	10.51 ±0.021	**6.24 ±0.031**	**0.77 ±0.026**	**1.86 ±0.017**	**6.26 ±0.014**
	KNN	19.11 ±0.021	18.85 ±0.013	24.47 ±0.011	18.94 ±0.022	19.16 ±0.021	24.50 ±0.015	16.15 ±0.012
	CNN	9.15 ±0.017	13.41 ±0.003	**6.57 ±0.001**	13.45 ±0.012	15.01 ±0.032	11.42 ±0.041	13.47 ±0.027
	LSTM	24.15 ±0.025	25.12 ±0.021	24.23 ±0.013	22.90 ±0.021	23.92 ±0.019	27.78 ±0.012	27.12 ±0.047
MCC	DST-CNN	**0.99 ±0.016**	**0.93 ±0.011**	0.88 ±0.015	**0.93 ±0.023**	**1.00 ±0.022**	**0.98 ±0.043**	**0.93 ±0.012**
	KNN	0.78 ±0.011	0.79 ±0.021	0.74 ±0.019	0.79 ±0.027	0.78 ±0.031	0.74 ±0.029	0.81 ±0.025
	CNN	0.90 ±0.016	0.85 ±0.012	**0.92 ±0.013**	0.85 ±0.014	0.83 ±0.020	0.87 ±0.017	0.85 ±0.021
	LSTM	0.74 ±0.011	0.73 ±0.013	0.74 ±0.017	0.76 ±0.011	0.75 ±0.014	0.69 ±0.012	0.69 ±0.023

Table 3
The key evaluation indexes of combined appliances for No. 2 building.

Index	Method	Refrig +WH	Refrig +Micro	Refrig +Inco	Refrig +WH +Micro	Refrig +WH +Inco	Refrig +Micro +Inco	Refrig +WH +Micro +Inco
ZL	DST-CNN	**0.73 ±0.081**	**0.32 ±0.075**	**2.47 ±0.064**	**5.32 ±0.052**	**0.99 ±0.047**	**1.01 ±0.052**	**2.31 ±0.037**
	KNN	25.96 ±0.037	26.61 ±0.051	25.21 ±0.031	23.91 ±0.035	23.87 ±0.029	26.10 ±0.024	23.89 ±0.041
	CNN	13.71 ±0.041	13.56 ±0.032	11.39 ±0.054	13.59 ±0.025	14.72 ±0.022	15.04 ±0.021	10.14 ±0.043
	LSTM	33.43 ±0.049	29.35 ±0.041	28.57 ±0.035	27.88 ±0.036	33.03 ±0.037	27.88 ±0.038	26.41 ±0.031
MCC	DST-CNN	**1.00 ±0.075**	**1.00 ±0.014**	**0.97 ±0.027**	**0.95 ±0.037**	**1.00 ±0.042**	**1.00 ±0.013**	**0.98 ±0.035**
	KNN	0.72 ±0.083	0.71 ±0.041	0.73 ±0.015	0.75 ±0.034	0.75 ±0.032	0.71 ±0.021	0.75 ±0.028
	CNN	0.85 ±0.131	0.85 ±0.111	0.87 ±0.105	0.85 ±0.087	0.84 ±0.073	0.83 ±0.042	0.88 ±0.065
	LSTM	0.63 ±0.111	0.65 ±0.017	0.66 ±0.125	0.69 ±0.097	0.64 ±0.067	0.69 ±0.034	0.71 ±0.054

methods. Tables 2 and 3 show the comparison results of the evaluation indexes with four methods in combined electrical appliance identification of two buildings (No. 1 building and No. 2 building). From the results, in multi-load identification, the DST-CNN method has the significant performance advantage.

Computation time is also an indicator to evaluate the performance of the model. In the paper, the training data set from the No.1 building includes 208901 samples are used to verify the training time of model. The training time on the data set of different sizes (20%–80%)is used. The training time of the different methods is presented in Fig. 14. When

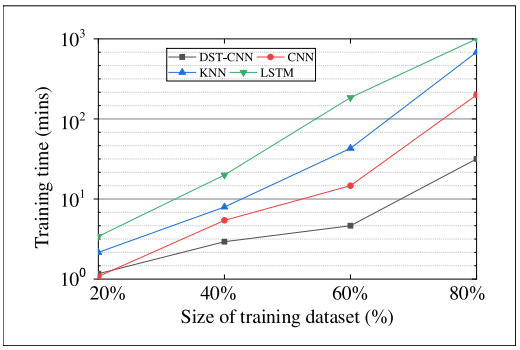

Fig. 14. Comparison of training time on the different size of data set.

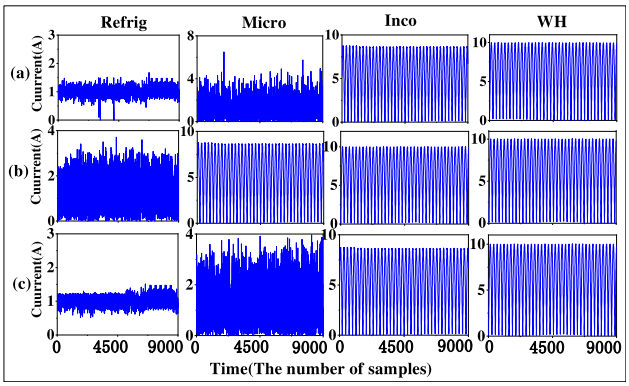

Fig. 16. The results of residential load identification in No. 3 family.

the size of the data set is only 20%, the training time of CNN model is the shortest. As the size of the data set increasing, the DST-CNN model is the fastest than the other models. However, the training time is also depends on the size of the training dataset and the performance of the computer. The performance of DST-CNN model and the outstanding characterization learning ability can improve the computation time.

4.4. Residential households load identification experiment

In this section, the performance and practicability of our method in residential households load identification are verified. We randomly sampled the real steady-state current signals in No. 3 building in any three days of May, and the sampling period is 24 h of one day in Fig. 15(a). From the figure, the main power consumption time of household equipment contains 6:00 a.m.-8:00 a.m., 11:00 a.m.-1:00 p.m., 6:00 p.m.-8:00 p.m., and 9:00 p.m.-10:30 p.m.. Considering the imaginable error in raw data set, the method of event detection and pre-processing are used, and used the DST-CNN method to realize the load identification. The frequency-domain characteristic map and the experimental results are showed in Fig. 15(b), Fig. 16, respectively.

From Fig. 16, the results of load identification for the four types of household appliance loads are clearly showed, and the each row represents a complete collection date cycle, which according to the three datasets in Fig. 15. Induction cooker and water heater work in the same pattern, refrigerator and micro-wave oven work in three different types work patterns respectively, and each work pattern present a different current characteristic waveform.

In the residential households loads identification experiment, a complete experimental process is clearly presented, and these results prove the practicability of the proposed method for NILMI. These results also demonstrate that our model can improve the ability of load abstract feature representation and the accuracy of the load fine-grained identification. Especially, in the multi-load identification experiment, the complex load features are accurately captured, which concludes that the complex state of load can be considered as a combination of the single state. Using the representative dataset to validate the proposed method, and it demonstrates that the method can improve the accuracy of load feature extraction and identification. Using the different data sets to validate the method, and it demonstrates the high reliability.

5. Conclusions

In this article, a reliable deep learning-based algorithm framework for IoT loads identification, DST-CNN, is proposed to address the problems of the incomplete discrimination of load feature, the accuracy and reliability of the identification mode, and the generalization and global optimization. DST-CNN method enhances the effectiveness of the load signal by AM-PCA method, improves the ability of load abstract characteristics representation by reconstructing DNN, and achieves the fine-grained accurate identification of load by a LSTM-RNN with the spatio-temporal feature recognition mechanism. The significant performance improvements of our proposed method are verified by extensive experiments. Applying the proposed method help to establish the hierarchical knowledge base of electrical equipment, which covering the information such as power, operating mode and time consumption. This way can help the demand-side of the smart grids to accurately obtain the power consumption information of the user, and increase the

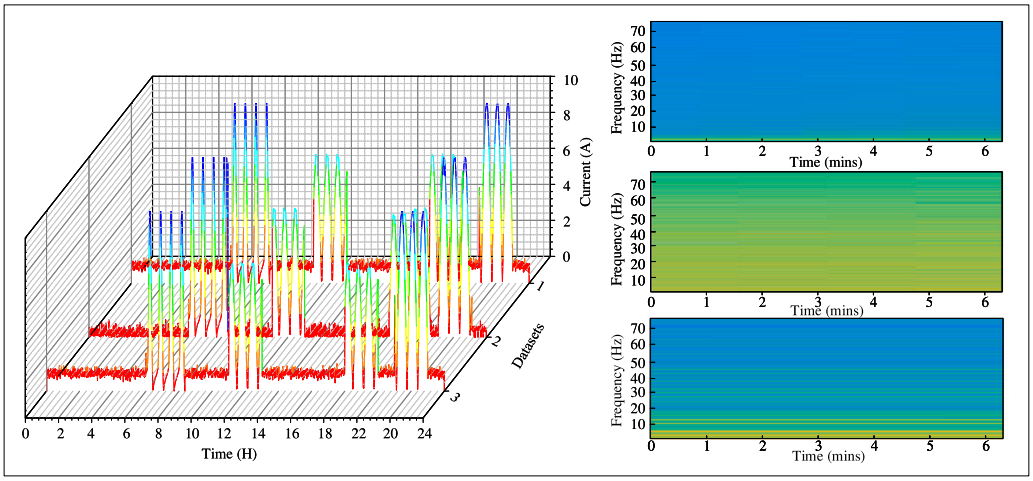

Fig. 15. The time-domain waveform and the frequency feature map.(a)The time-domain waveform.(b)The frequency feature mapping.

electric power efficiency of the customers and load shifting in IoT load monitoring system. In our future work, DST-CNN will be used on the more complex data sets, without reducing the accuracy of recognition and adding the time complexity.

Declaration of competing interest

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

Acknowledgment

The authors of this paper were supported by S&T Program of Hebei through grant 20310101D.

References

[1] T. Wang, M.Z.A. Bhuiyan, G. Wang, M.A. Rahman, J. Wu, J. Cao, Big data reduction for a smart city's critical infrastructural health monitoring, IEEE Commun. Mag. 56 (3) (2018) 128–133.
[2] D. Garcia-Perez, D. Perez-Lopez, I. Diaz-Blanco, A. Gonzalez-Muniz, A.A. Cuadrado-Vega, Fully-convolutional denoising auto-encoders for NILM in large non-residential buildings, IEEE Trans. Smart Grid PP (99) (2020) 1.
[3] R. Jones, A. Rodriguez-Silva, S. Makonin, Increasing the accuracy and speed of universal non-intrusive load monitoring (UNILM) using a novel real- time steady-state block filter, in: Proceedings of the IEEE Power & Energy Society Innovative Smart Grid Technologies Conference, ISGT, 2020.
[4] H. Imen, N. Etinkaya, J.C. Vasquez, J.M. Guerrero, A microgrid energy management system based on non-intrusive load monitoring via multitask learning, IEEE Trans. Smart Grid 12 (2) (2020).
[5] S. Wang, L. Du, J. Ye, D. Zhao, Deep generative model for non-intrusive identification of EV charging profiles, IEEE Trans. Smart Grid PP (99) (2020) 1.
[6] A. Faustine, L. Pereira, C. Klemenjak, Adaptive weighted recurrence graphs for appliance recognition in non-intrusive load monitoring, IEEE Trans. Smart Grid PP (99) (2020) 1.
[7] H. Tao, M.Z.A. Bhuiyan, M.A. Rahman, T. Wang, J. Wu, a. Salih, T. Hayajneh, Trustdata: Trustworthy and secured data collection for event detection in industrial cyber-physical system, IEEE Trans. Ind. Inf. 16 (5) (2020) 3311–3321.
[8] L. Pereira, N. Nunes, An empirical exploration of performance metrics for event detection algorithms in non-intrusive load monitoring, Sustainable Cities Soc. (2020) 102399.
[9] G. Rajendiran, M. Kumar, C. Joshua, K. Srinivas, Energy management using non-intrusive load monitoring techniques - state-of-the-art and future research directions, Sustainable Cities Soc. 62 (2020) 102411.
[10] W. Kong, Z.Y. Dong, B. Wang, J. Zhao, J. Huang, A practical solution for non-intrusive type II load monitoring based on deep learning and post-processing, IEEE Trans. Smart Grid PP (99) (2019) 1.
[11] M.A. Mengistu, A.A. Girmay, C. Camarda, A. Acquaviva, E. Patti, A cloud-based on-line disaggregation algorithm for home appliance loads, IEEE Trans. Smart Grid (2018) 1.
[12] J. Wang, G. Geng, K.L. Chen, H. Liang, W. Xu, Event-based non-intrusive home current measurement using sensor array, IEEE Trans. Smart Grid (2017) 1.
[13] N. Henao, K. Agbossou, S. Kelouwani, Y. Dube, M. Fournier, Approach in nonintrusive type I load monitoring using subtractive clustering, IEEE Trans. Smart Grid PP (2) (2015) 1.
[14] S. Singh, A. Majumdar, Non-intrusive load monitoring via multi-label sparse representation-based classification, IEEE Trans. Smart Grid PP (99) (2019) 1.
[15] S. Pascal A., M. Iosif, Statistical and electrical features evaluation for electrical appliances energy disaggregation, Sustainability 11 (3222) (2019).
[16] Bonfigli, Roberto, Felicetti, Andrea, Principi, Emanuele, Fagiani, Marco, Squartini, Stefano, Denoising autoencoders for non-intrusive load monitoring: Improvements and comparative evaluation, Energy Build. (2018).
[17] X. Wu, D. Jiao, Y. Du, Automatic implementation of a self-adaption non-intrusive load monitoring method based on the convolutional neural network, Processes (2020).
[18] D. Zhao, C. Chen, Z. Fang, Non-intrusive appliance identification with appliance-specific networks, IEEE Trans. Ind. Appl. (2020).
[19] R. Gopinath, M. Kumar, K. Srinivas, Feature mapping based deep neural networks for non-intrusive load monitoring of similar appliances in buildings, in: Proceedings of the in The 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, BuildSys '20, 2020.
[20] S. Mohamad, H. Bouchachia, Deep online hierarchical dynamic unsupervised learning for pattern mining from utility usage data, Neurocomputing (2019).
[21] Q. Zhao, Y. Xu, Z. Wei, Y. Han, Non-intrusive load monitoring based on deep pairwise-supervised hashing to detect unidentified appliances, Processes 9 (3) (2021) 505.
[22] F. Ciancetta, G. Bucci, E. Fiorucci, S. Mari, A. Fioravanti, A new convolutional neural network-based system for NILM applications, IEEE Trans. Instrum. Meas. PP (99) (2020) 1.
[23] Y. Tian, H. Wang, A. Li, S. Shi, J. Wu, Non-intrusive load monitoring using inception structure deep learning, in: 2020 10th International Conference on Power and Energy Systems, ICPES, 2020.
[24] K. Jihyun, T. Le, K. Howon, Nonintrusive load monitoring based on advanced deep learning and novel signature, Comput. Intell. Neurosci. 2017 (2017) 4216281.
[25] K. He, L. Stankovic, L. Jing, V. Stankovic, Non-intrusive load disaggregation using graph signal processing, IEEE Trans. Smart Grid 9 (3) (2018) 1739–1747.
[26] S. Singh, A. Majumdar, Non-intrusive load monitoring via multi-label sparse representation-based classification, IEEE Trans. Smart Grid PP (99) (2019) 1.
[27] M. Qureshi, C. Ghiaus, N. Ahmad, A blind event-based learning algorithm for non-intrusive load disaggregation, Int. J. Electr. Power Energy Syst. 129 (1) (2021) 106834.
[28] Q. Liu, K.M. Kamoto, X. Liu, M. Sun, N. Linge, Low-complexity non-intrusive load monitoring using unsupervised learning and generalized appliance models, IEEE Trans. Consum. Electron. 65 (1) (2019) 28–37.
[29] N. Chumrit, C. Weangwan, N. Aunsri, ECG-based Arrhythmia detection using average energy and zero-crossing features with support vector machine, in: 2020-5th International Conference on Information Technology, InCIT, 2021.
[30] D.S. Yang, X.T. Gao, L. Kong, Pang, An event-driven convolutional neural architecture for non-intrusive load monitoring of residential appliance, IEEE Trans. Consum. Electron. 66 (2) (2020) 137–182.
[31] JiangJie, KongQiuqiang, D. Plumbleymark, GilbertNigel, HoogendoornMark, M. Roijersdiederik, Deep learning-based energy disaggregation and on/off detection of household appliances, ACM Trans. Knowl. Discov. Data (2021).
[32] S. Houidi, D. Fourer, F. Auger, On the use of concentrated time-frequency representations as input to a deep convolutional neural network: Application to non intrusive load monitoring, Entropy 22 (9) (2020).
[33] J. Zhang, X. Chen, W. Ng, C.S. Lai, L.L. Lai, New appliance detection for non-intrusive load monitoring, IEEE Trans. Ind. Inf. PP (99) (2019) 1.
[34] A. Wang, Longjun, B. Chen, C. Xiaomin, Gang, D. Hua, Non-intrusive load monitoring algorithm based on features of V-I trajectory, Electr. Power Syst. Res. 157 (Apr.) (2018) 134–144.
[35] N. Miao, S. Zhao, Q. Shi, R. Zhang, Non-intrusive load disaggregation using semi-supervised learning method, in: Proceedings of the 2019 International Conference on Security, Pattern Analysis, and Cybernetics, SPAC, 2020.
[36] Y.H. Lin, A parallel evolutionary computing-embodied artificial neural network applied to non-intrusive load monitoring for demand-side management in a smart home: Towards deep learning, Sensors 20 (6) (2020) 1649.
[37] G. Cui, B. Liu, W. Luan, et al., Estimation of target appliance electricity consumption using background filtering, IEEE Trans. Smart Grid 1 (2019).
[38] B. Hs, A. Sm, B. Mt, Unsupervised Bayesian non parametric approach for non-intrusive load monitoring based on time of usage - sciencedirect, Neurocomputing (2021).
[39] D. Su, Q. Shi, H. Xu, W. Wang, Nonintrusive load monitoring based on complementary features of spurious emissions, ELECTRONICS (2019).
[40] M. Xia, W. Liu, K. Wang, W. Song, C. Chen, Y. Li, Non-intrusive load disaggregation based on composite deep long short-term memory network, J. Magn. Mater. Dev. 160 (113669) (2020).
[41] M.C. Pegalajar, L. Ruiz, M. Cuéllar, R. Rueda, Analysis and enhanced prediction of the Spanish electricity network through big data and machine learning techniques, Internat. J. Approx. Reason. 133 (11) (2021).
[42] G. M. Junaid, U. Gul Malik, P. Anand, M. Jihoon, R. Seungmin, Mid-term electricity load prediction using CNN and Bi-LSTM, J. Supercomput. (2021).
[43] H. Zhou, D. Ren, H. Xia, M. Fan, H. Huang, AST-GNN: AN attention-based spatio-temporal graph neural network for interaction-aware pedestrian trajectory prediction, Neurocomputing (2021).
[44] X. Yin, G. Wu, J. Wei, Y. Shen, B. Yin, Multi-stage attention spatial-temporal graph networks for traffic prediction, Neurocomputing 428 (6) (2020).
[45] E. Luo, M.Z.A. Bhuiyan, G. Wang, M.A. Rahman, J. Wu, M. Atiquzzaman, Privacyprotector: Privacy-protected patient data collection in IoT-based healthcare systems, IEEE Commun. Mag. 56 (2) (2018) 163–168.
[46] A. Faustine, L. Pereira, Improved appliance classification in non-intrusive load monitoring using weighted recurrence graph and convolutional neural networks, Energies 13 (2020).
[47] Zain, J. Mohamad, Hayajneh, Thaier, Abdalla, N. Ahmed, Hassan, M. Mehedi, Bhuiyan, M.Z.a. Alam, Secured data collection with hardware-based ciphers for IoT-based healthcare, IEEE Internet Things J. (2019).
[48] J.A. Cadzow, Signal enhancement-a composite property mapping algorithm, Acoust Speech Signal Process IEEE Trans 36 (1) (1988) 49–62.
[49] X. Peng, M. Bergsneider, H. Xiao, Pulse onset detection using neighbor pulse-based signal enhancement, Med. Eng. Phy. 31 (3) (2009) 337–345.
[50] M. Lin, M. Tsai, Development of an improved time–frequency analysis-based nonintrusive load monitor for load demand identification, IEEE Trans. Instrum. Meas. 63 (6) (2014) 1470–1483.
[51] K. Jihyun, L. Thi-Thu-Huong, K. Howon, Nonintrusive load monitoring based on advanced deep learning and novel signature, Comput. Intell. Neurosci. 2017 (2017) 4216281.
[52] H. Peng, J. Li, Q. Gong, Y. Ning, L. He, Motif-matching based subgraph-level attentional convolutional network for graph classification, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, No. 4, 2020, pp. 5387–5394.

Yanmei Jiang is currently working toward the Ph.D. degree at State Key Laboratory of Reliability and Intelligence of Electrical Equipment in Hebei University of Technology. His research interests include machine learning, application of intelligent power equipment,Internet of Things reliability.

Mingsheng Liu is currently a professorat State Key Laboratory of Reliability and Intelligence of Electrical Equipment in Hebei University of Technology. His research focuses on network and information security, computer network technology and application, big data.

Hao Peng is currently an Assistant Professor at the School of Cyber Science and Technology, and Beijing Advanced Innovation Center for Big Data and Brain Computing in Beihang University. His research interests include representation learning, machine learning and graph mining.

Md Zakirul Alam Bhuiyan(Senior Member, IEEE) was as an Assistant Professor with Temple University and a Postdoctoral Research Fellow with the School of Information Science and Engineering, Central South University. He is currently an Assistant Professor with the Department of Computer and Information Sciences, Fordham University, NY, USA. His research focuses on dependability, cybersecurity, big data, and cyber physical systems. He is also a member of ACM.