



## OPEN ACCESS

## EDITED BY

Xin Ning,  
Institute of Semiconductors (CAS), China

## REVIEWED BY

Yassine Himeur,  
Qatar University, Qatar  
Bing Yan,  
Rochester Institute of Technology,  
United States  
Wang Wei,  
Chiba University, Japan

## \*CORRESPONDENCE

Min Xia,  
✉ xiamin@nuist.edu.cn

## SPECIALTY SECTION

This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 09 January 2023

ACCEPTED 08 February 2023

PUBLISHED 20 February 2023

## CITATION

Yu F, Wang Z, Zhang X and Xia M (2023),  
DRA-net: A new deep learning framework  
for non-intrusive load disaggregation.  
*Front. Energy Res.* 11:1140685.  
doi: 10.3389/fenrg.2023.1140685

## COPYRIGHT

© 2023 Yu, Wang, Zhang and Xia. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# DRA-net: A new deep learning framework for non-intrusive load disaggregation

Fang Yu<sup>1</sup>, Zhihua Wang<sup>2</sup>, Xiaodong Zhang<sup>3</sup> and Min Xia<sup>3\*</sup>

<sup>1</sup>China Electric Power Research Institute, Nanjing, China, <sup>2</sup>State Grid Shanghai Municipal Electric Power Company, Shanghai, China, <sup>3</sup>Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing University of Information Science and Technology, Nanjing, China

The non-intrusive load decomposition method helps users understand the current situation of electricity consumption and reduce energy consumption. Traditional methods based on deep learning are difficult to identify low usage appliances, and are prone to model degradation leading to insufficient classification capacity. To solve this problem, this paper proposes a dilated residual aggregation network to achieve non-intrusive load decomposition. First, the original power data is processed by difference to enhance the data expression ability. Secondly, the residual structure and dilated convolution are combined to realize the cross layer transmission of load characteristic information, and capture more long sequence content. Then, the feature enhancement module is proposed to recalibrate the local feature mapping, so as to enhance the learning ability of its own network for subtle features. Compared to traditional network models, the null-residual aggregated convolutional network model has the advantages of strong learning capability for fine load features and good generalisation performance, improving the accuracy of load decomposition. The experimental results on several datasets show that the network model has good generalization performance and improves the recognition accuracy of low usage appliances.

## KEYWORDS

non-intrusive, load disaggregation, deep learning, feature extraction, energy efficiency, residential electricity

## 1 Introduction

With the development trend of smart grid, the traditional intrusive load monitoring method has many problems, such as high construction cost and difficult application, which makes the non-invasive load monitoring method a unique way to solve these problems. Non-intrusive load decomposition can help power companies more easily obtain the power consumption of users and understand the power consumption of various electrical appliances. Non-intrusive load decomposition can more accurately predict the distribution of residential power consumption and the total amount of residential load by providing the power consumption of each electrical appliance of users, reduce planning investment, save budget, and avoid unnecessary waste of power resources. It is also conducive to the scientific formulation of relevant policies for dynamic demand response by power companies, the adjustment of electricity prices, the evaluation of relevant projects and the more reasonable allocation of power resources, and the formation of a more benign and friendly interaction between users and power companies, so as to achieve the results of peak shaving and valley

filling, mutual benefit and win-win results. On the other hand, if a family can know more details about electricity consumption, it will consciously reduce energy waste. For example, most American and British families install smart meters in their homes to facilitate users to learn about the low peak electricity price information in time, promote users to use electricity at night or at low prices, alleviate the pressure of peak electricity consumption, avoid power loss, and indirectly improve the economic benefits of power sources (Zhao et al., 2019). In 1992, Hart proposed non-invasive load monitoring (NILM). Its essence is non-invasive load decomposition NILD (Hart, 1992), that is, the total energy consumption is decomposed into a single device to analyze the electricity consumption behavior of residential users. This provides effective feedback on residential electricity consumption, helping users save energy and reduce electricity charges (Paterakis et al., 2017). Hart's method is mainly to extract steady-state features for power decomposition. Based on Hart's algorithm, a simple non-invasive charge load monitoring system can be designed. However, this algorithm can only be used for a small number of electrical appliances, and the number of types of features extracted is small. When there are many types and numbers of electrical appliances, the decomposition accuracy of this algorithm decreases significantly (Dash and Sahoo, 2022).

In view of the problems in the above non-invasive load decomposition, Inagaki et al. (2011) and others used the integer programming method to monitor the load of household power equipment, but it is only applicable to equipment in discrete operation mode. Kolter et al. (2010) studied sparse coding algorithm to improve decomposition performance, but this method is only applicable to data sets containing low resolution data types. Lin Y. and Tsai (2014) and Chang et al. (2013) used particle swarm optimization algorithm to carry out non-invasive charge load decomposition experiments for a small number of several electrical appliances. This algorithm can decompose the total power data to each electrical equipment at the same time, but the error of the decomposition results is still large. The optimization method is based on load characteristic analysis. First, the static and dynamic characteristics of the load should be modeled. The total load curve is the superposition of multiple loads. The objective of optimization is to obtain the optimal load coefficient (i.e., the contribution of each load), so as to minimize the residual between the superimposed total load and the actual load. Piga et al. (2016) proposed a sparse optimization algorithm for non-invasive charge load decomposition, which reduced the decomposition error to a certain extent. Ahmadi and Marti (2015) proposed a non-invasive charge load decomposition experiment based on feature matching (also called load information matching), which effectively solved the problem of high similarity between load features. Johnson and Willisky (2013); Luan et al. (2022); Xia et al. (2021) used the Hidden Markov Model to perform non-invasive charge load decomposition. Similar to the combination optimization algorithm, these algorithms first obtain the state power of electrical appliances through clustering. Its encoding and decoding process is the process of optimizing the power values obtained by these clustering, and the decomposition results are also the combination of power values obtained by clustering, which cannot obtain more accurate electrical power consumption values (Himeur et al., 2020a; Fan et al., 2021). Compared with

hidden Markov algorithm, Tsai and Lin (2012) proposed a method achieves more accurate non-invasive load decomposition through K-nearest neighbor regression algorithm. However, when the power consumption difference between appliances is large, this algorithm cannot achieve accurate decomposition. Other researches, such as algorithms based on Adaboost algorithm (Hassan et al., 2014), 2D phase encoding algorithm (Himeur et al., 2020b; Himeur et al., 2021a), fuzzy algorithm (Lin Y. H and Tsai, 2014), bagging tree algorithm (Himeur et al., 2020c), histogramming descriptor algorithm (Himeur et al., 2021b) and neural network algorithm, have made certain achievements in non-invasive charge load decomposition tasks. For the machine learning model, non-intrusive load decomposition is to take the total power time series data as the input, take the power data of each electrical appliance as the output (fitting method) or take the electrical appliance category as the output (classification). These two tasks are consistent in nature, although it is more difficult to take the power data of each electrical appliance as the output. For machine learning methods, a large number of sample data are needed for training. Non-intrusive load decomposition based on machine learning is to fit the machine learning model through training of a large number of samples, so as to obtain the power distribution of various loads under different total power conditions, which has no essential difference from traditional machine learning fitting and classification. For 1/60 Hz sampling data, it can meet the needs of non-intrusive load decomposition very well. In Lin's work, the fuzzy C-means algorithm based on particle swarm optimization is combined with the fuzzy neural algorithm for non-invasive charge load decomposition experiments. This algorithm can identify the state of an electrical appliance at a certain time, and also solve the problem of high similarity between the power consumption characteristics of electrical appliances. Park et al. (2019) proposed an equipment status recognition algorithm based on neural network, which is simple and fast in decomposition. Welikala et al. (2019) proposed a NILD method, which combines the application usage patterns (AUPs) of equipment to improve the state recognition performance of high-frequency appliances. Himeur et al. (2021c) proposed a histogram post-processing of 2D local binary patterns for smart grid applications. Guo et al. (2021) proposed a multi-model combination model for non-intrusive load disaggregation, which It can integrate the advantages of various methods and improve the accuracy of decomposition. The traditional methods mainly use artificial features to realize the identification of electrical appliances by optimizing methods, but it is difficult to extract effective artificial features. The algorithm is highly sensitive to noise and has low decomposition accuracy.

Traditional load feature extraction needs manual design, so it is difficult to extract effective features, and it is difficult to analyze the features of time series. Recently, deep learning technology has been widely used in various fields (Qu et al., 2021; Chen et al., 2022; Gao et al., 2022; Song et al., 2023). The application of deep learning in non-invasive load decomposition has gradually attracted researchers' attention. So far, there are a large number of load record data for model training. For model training, we need to record the power data of a household user and each load. At present, there are many data sets, such as UK-DALE public data set and WikiEnergy data set, which provides a data basis for the application of deep learning. Different from traditional pattern recognition,

deep learning can automatically extract features without manual extraction. In Kelly's experiment (Kelly and Knottenbelt, 2015), it is proved that AutoEncoder method has the best effect in sequence pair sequence method. Singh and Majumdar (2018), Singh and Majumdar (2019) proposed deep sparse coding for non-intrusive load monitoring, improved decomposition efficiency. Xia et al. (2019) constructed a deep dilated residual network for load feature extraction, it can improve feature utilization. Jia et al. (2021) used bidirectional dilated residual network to realize the sequence to point non-intrusive load decomposition. At present, there are two main load decomposition methods: sequence to sequence and sequence to point. Sequence-to-sequence refers to the direct decomposition of the input sequence into different load sequences. Instead of training a network to predict a window, the sequence to point method is only to predict the midpoint element of input window. The idea is that the input of the network is a mains window, the output is the power at one point of various electrical devices. Zhang et al. (2016) realized sequence to sequence and sequence to point non-invasive charge load decomposition using convolutional neural network. In Zhang's experiment, sequence to point decomposition method has achieved good results in the decomposition of most electrical equipment. Compared with other deep learning methods, convolutional neural network has been proved to be more effective in the application of non-invasive load decomposition. However, Zhang's experiment uses a relatively shallow convolutional neural network, which is prone to the phenomenon of gradient disappearance, and cannot extract the deep level charge load characteristics. It is difficult to capture the relationship between long time series data. Xia et al. (2020) constructed a deep LSTM model to realize the decomposition of sequences into multiple sequences, and improved the decomposition accuracy through depth feature extraction. However, the current deep learning model is prone to model degradation and other problems, resulting in insufficient fitting ability. In addition, because it is difficult to extract the features of low usage appliances, the weight of low usage appliances in deep learning training is too small. Therefore, the existing deep learning model has poor decomposition effect on low utilization rate appliances.

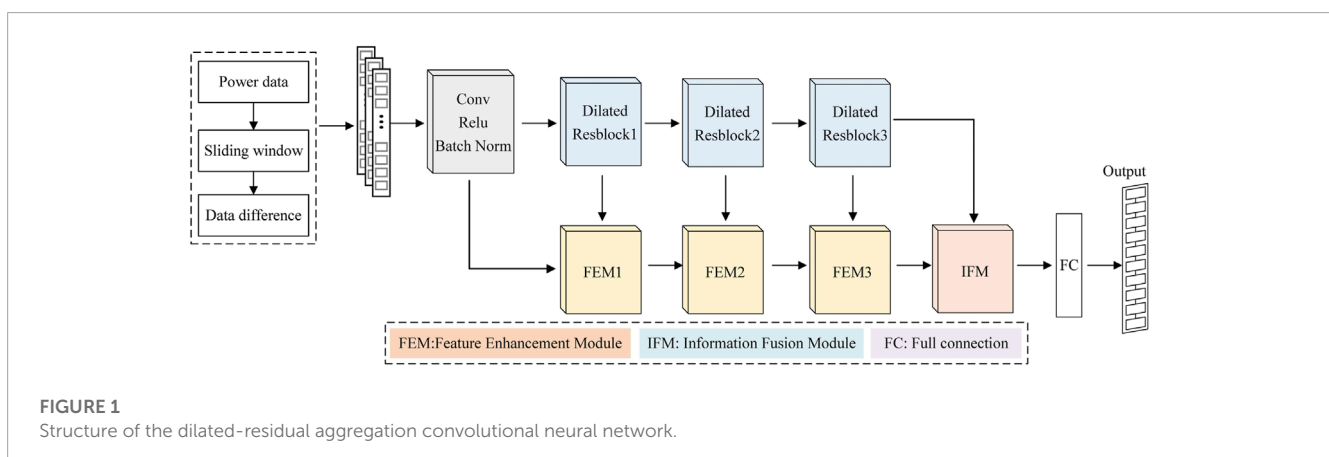
To solve above problem, a dilated-residual aggregation network (DRA-Net) is constructed and applied to non-invasive load disaggregation. The network model increases the receptive field

of convolution kernel through hole convolution to capture more features. In addition, a feature enhancement module is proposed to improve the learning ability of the model to fine load features, and further improve the generalization performance of non-invasive load decomposition. In conclusion, our contributions are as follows: 1) The differential processing of raw power data enhances the ability of data expression. 2) A structure combining residual structure and dilated convolution is proposed to realize cross-layer transmission of load characteristic information and capture more long sequence content. 3) A feature enhancement module is proposed to recalibrate the local feature mapping to enhance the learning ability of network for fine features.

## 2 Dilated-residual aggregation convolutional neural network

Since different load devices in residential houses have different electrical characteristics and some other influencing factors such as interference noise, this work realizes the optimization and improvement of the common convolutional network structure, and proposes a new network model, Dilated-residual Aggregation Convolutional Neural Network (DRANet) to realize Non-intrusive power load disaggregation, whose overall structure is shown in Figure 1.

As shown in Figure 1, the network model includes the ordinary convolutional layer, the Dilated Resblock, Feature Enhancement Module (FEM), Information Fusion Module (IFM), and Fully Connected Layer (FC). There are a total of three void residual modules, namely, Dilated Resblock1, Dilated Resblock2, and Dilated Resblock3. There are a total of three feature enhancement modules, namely, FEM1, FEM2, and FEM3. The total load power is differentially processed and then recombined with the original data as the input side, which enriches the edge information of residential power data. The convolutional layer of the Dilated-residual Aggregation Convolutional Neural Network enhances the extraction of load features of different residential electrical devices by combining multiple convolutional kernels to retain the basic load characteristic information. The initial convolutional layer for load disaggregation feature mapping is followed by three void residual modules. The non-intrusive power load disaggregation



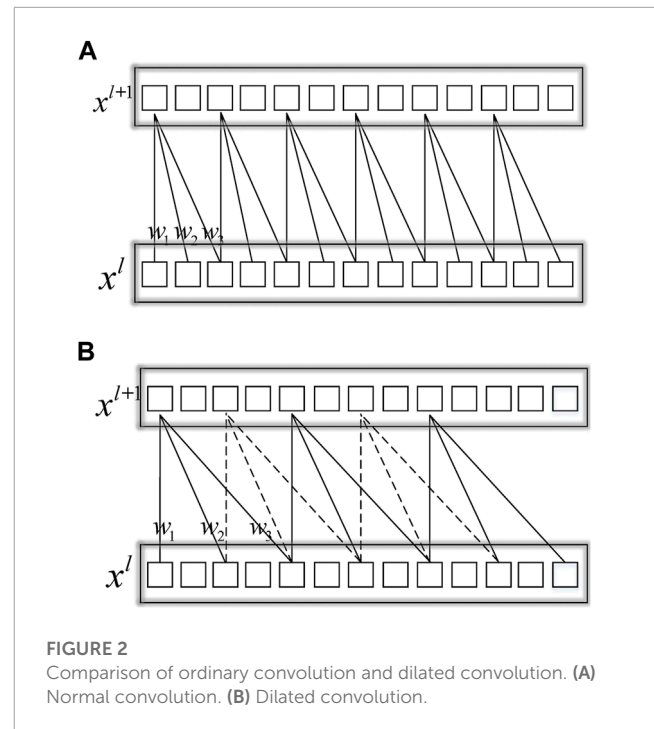
task enhances the model's ability to fit the load features without increasing the number of parameters through these three cavity residual modules. At the same time, it ensures that the signal passed layer by layer in the network will not be lost when back propagation is performed. The dilated residual module is based on the idea of "cross-layer connection" and uses the residual connection to further extract features from the low-order load features Hu et al. (2022). The higher-order residential load feature mapping contains more abstract load characteristics and timing information.

To enhance the disaggregation performance of the network for load, a feature enhancement module is proposed to process the output mapping of different stages in the network to obtain the attention weight matrix. This matrix is multiplied and summed with the corresponding vectors with the output of the dilated residual module. This module facilitates the integration and strengthening of the base load characteristics extracted by the first three void residual modules, and fully utilizes the load characteristics and timing information of each network stage, which makes the fitting ability of the whole network structure enhanced, especially for the load devices with low usage frequency like washing machines and dishwashers, and has a great disaggregation improvement effect. The output dilated residual convolutional feature mapping and the feature-enhanced load feature mapping both contain a large amount of different higher-order load information. Using the dimensional splicing in the information fusion module and conventional convolutional operation processing, these two parts are then integrated, and finally the power prediction results of the electrical equipment are output by the full-connected operation. Dilated-residual Aggregation Convolutional Neural Network constructed in this section is mainly designed to make up for the defects of insufficient utilization of residential load characteristics, poor disaggregation of low-use appliances and disappearance of gradients, reflecting the more excellent feature extraction capability and learning ability of the dilated residual aggregation convolutional network.

### 2.1 Dilated resblock

The dilated convolution uses the parameter Dilated Factor (DF) to adjust the size of the dilated convolution (Miao et al., 2022). Since loads like washing machines and dishwashers are used less frequently and have sparse temporal features, the proposed dilated convolution allows resampling of the underlying load feature mapping. Pooling and down-sampling operations cause the loss of temporal information of the load, while the advantage of the dilated convolution is that it can both replace the pooling effect and increase the field of sensation exponentially (field of sensation refers to the corresponding size of the convolution kernel, that is, the range of the convolution for the load series.), allowing each convolution output to capture a larger range of feature information, which has a good feature extraction effect for load appliances like washing machines without adding extra redundant number of parameters.

The dilated convolution kernel in the Dilated-residual Aggregation Convolutional Neural Network is shown in Figure 2B. Figure 2A shows the ordinary convolutional kernel convolution, where  $x_l$  and  $x_{l+1}$  are the input and output of  $l + 1$  layer respectively. Assuming that the convolutional kernel size  $kernel$  is a and the



step *stride* is 1, when the hole rate  $d$  is 1, i.e., the number of filled "0" weights is 0. From the calculation of Eq. 1, we can see that the mapping range length of layer  $l + 1$  ( $L_{l+1}$ ) and layer  $l$  ( $L_l$ ) is the same.

$$L_{l+1} = \frac{[L_l + 2 \times padding - kernel - (kernel - 1) \times (d - 1)]}{stride} + 1, \quad (1)$$

where *padding* is the number of padding zeros. The dilated rate  $d$  is 2 in Figure 2B, and the perceptual field is expanded to  $5 \times 1$ . So, the advantage of the dilated convolution lies in the ability to increase the local receptive field during the convolution operation and capture more information about the load characteristics without introducing additional parameters. For load feature extraction, dilated convolution can control the receptive field without changing the size of the feature map, so as to extract multi-scale information and effectively improve the accuracy of load decomposition. The combination of the three-layer dilated convolution and the residual connection constitutes a dilated residual module, as shown in Figure 3.

The feature mapping is performed sequentially using the dilated convolution kernel with convolution of the dilated rate of 1, 2, and 3, including Leaky-Relu (Leaky ReLU is the commonly used activation function of convolution neural network at present), Batch Normalization (Similar to common data standardization, it is a way to unify scattered data and a common method to optimize neural network at present.), and other operations for processing, and the reason why Relu is not used as a non-linear activation function is that when the input value of the convolution layer is negative, the learning speed of Relu will be slow (Wang et al., 2022), even deactivating the neurons and preventing them from updating weights, resulting in the disappearance of the network gradient. Leaky-Relu activation function can correct the distribution of load data and retain the negative values in the gradient calculation process (Lu et al., 2022), which indirectly improves the

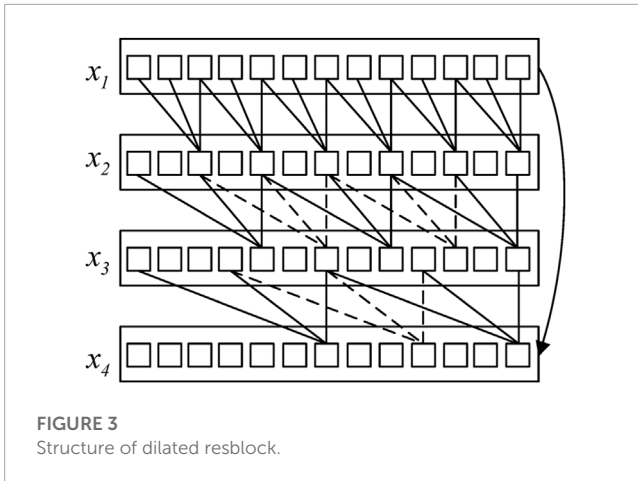


FIGURE 3 Structure of dilated resblock.

retention of power load timing information in the network. The residual connection also solves the problem of network degradation caused by the gradual disappearance of information in the reverse transmission of network layers, so the dilated residual module enhances the fitting ability of the network itself to the load samples and improves the disaggregation accuracy of the model.

## 2.2 Feature enhancement module

The structure of the feature enhancement module proposed in this work is shown in Figure 4. The two input mappings of the internal structure of the null residual aggregated convolutional network are used as the input side of the feature reinforcement module, which are feature map1 and feature map2. Where feature map1 is the output feature map of the current null residual module, while feature map2 is the output feature map of the previous stage. As known from Figure 4, there are two optimization branches of the feature reinforcement module, the first branch includes operations such as Convolution, GlobalAvgPooling (GAP), FC, Sigmoid, Reshape, etc. The second branch contains operations such as tensor multiplication of feature maps, i.e., feature rescaling. Feature map1 and feature map2 are processed by the feature reinforcement module to obtain the output mapping feature map3.

The specific process of the enhancement module in loading information is shown in Figure 5. In Figure 5, feature map2 is the previous stage feature mapping, so it is not consistent with the dimensionality of the output mapping feature map1 of the null residual module, and the dimensionality of feature map2 is C. The number of convolution kernels of the convolution is  $1 \times 1 \times C$ . The convolution operation of this  $1 \times 1$  convolution makes the two dimensions consistent, which is convenient for subsequent processing. The weight vector1 (weight vector1 in the figure) is obtained after the GlobalAvgPooling layer operation, which compresses the pre-multi-dimensional load feature map to a one-dimensional feature map. Weight vector1 is essentially a one-dimensional vector containing low-order load feature information, which characterizes the global information on the feature layer, and its dimension is  $1 \times C$ . weight vector1 is then processed by FC layer, Relu non-linear activation function to obtain weight vector2 (weight

vector2 in the figure), weight vector2 is a high-dimensional vector of higher-order global features obtained on the basis of weight vector1, whose dimension is also  $1 \times C$ . Through this series of operations, the weight vector2 further represents the change in dimensional response of residential housing load characteristics as described as follows:

$$W_1 = GAP(conv(F_1)), \tag{2}$$

$$W_2 = FC(ReLu(W_1)), \tag{3}$$

Where  $F_1$  is the feature mapping of feature map1,  $GAP(\cdot)$  is global average pooling,  $FC(\cdot)$  is fully connection.  $W_1$  and  $W_2$  is the weight vector 1 and weight vector 2, respectively.

Sigmoid non-linear activation processing is similar to a gated filtering mechanism to achieve a filtering function on load information. Sigmoid processes each feature of the weight vector2 to generate a different weight variable. When the load feature of a channel is more effective, its corresponding weight variable is closer to 1; when the load feature of a channel is invalid, its corresponding weight variable is closer to 0. In this way the feature reinforcement module filters the useless information effectively. Then after Reshape operation to complete the dimensional change, the weight vector2 processed by the above operation is multiplied with feature map2 to complete the feature rescaling of feature map2, and finally the reinforced output mapping feature map3 is obtained, the specific calculation is as Eq. 4.

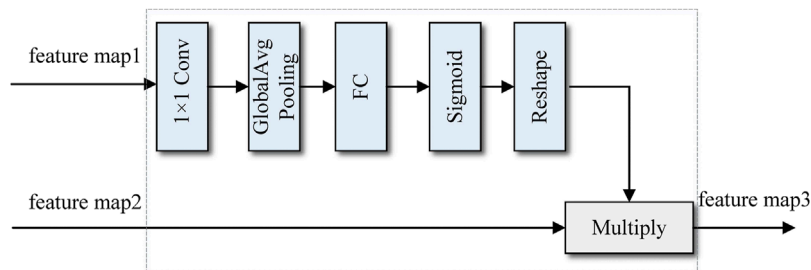
$$F_3 = F_2 \cdot Sigmoid(W_2), \tag{4}$$

Where  $F_3$  is the feature mapping of feature map3,  $Sigmoid(\cdot)$  is the Sigmoid function. The main role of the feature enhancement module is to integrate and optimize the features of the output mapping feature map1 of the dilated residual module and the output load feature mapping feature map2 of the previous part to complete the rescaling of the weights. This module makes use of the feature maps obtained at different stages of the network model, thus enabling the discrimination of the importance of the load features. The feature reinforcement module of the null residual aggregated convolutional network learns the dependencies of load timing information through the weight vector weight vector1 and weight vector weight vector2 and learns the importance of each load feature in the network accurately. Feature mappings that are favorable to the load decomposition task are given biased weight vectors, which serve to improve the decomposition accuracy, and conversely are given biased weight vectors that are small, which suppress irrelevant load features and achieve filtering of invalid information.

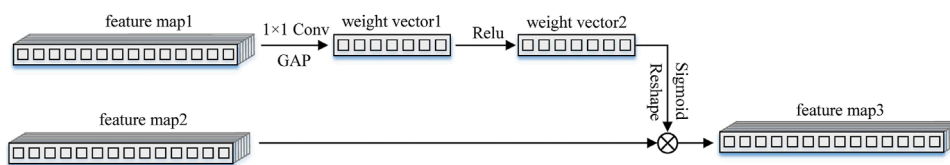
In summary, the null residual aggregated convolutional network model improves the decomposition accuracy of low-use appliances by completing the rescaling operation of load features through the feature reinforcement module.

## 2.3 Information fusion module

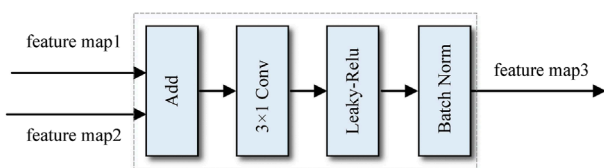
The structure of the information fusion module proposed in this work is shown in Figure 6. In total, it is composed of Add



**FIGURE 4**  
Structure diagram of feature enhancement module.



**FIGURE 5**  
Schematic diagram of the process of the feature enhancement module.



**FIGURE 6**  
Information fusion module structure diagram.

operation,  $3 \times 1$  Conv, Leaky-ReLU non-linear activation function, and Batch Normalization. The output feature mapping of the null residual module is different from the output feature mapping of the feature reinforcement module, so the two parts are fused using the information fusion module. The dimensionality of feature map1 and feature map2 of Figure 6 is made consistent using the base convolution, and the dual input feature mappings are fused by Add operation, and feature re-extraction is performed using  $3 \times 1$  convolution. After that, the load feature mapping is further adjusted and optimized using Leaky-ReLU activation function, Batch Normalization and other operations. Finally the output feature mapping of the information fusion module is a one-dimensional vector, and then the prediction results of residential load devices are obtained by the full join operation.

The null residual aggregated convolutional network uses multiple null residual blocks to extract the electric load features of residential houses, which improves the network model's ability to encode and decode load information, and also uses residual connections to transfer feature information to further ensure the effectiveness of load transfer between layers and avoid problems

such as gradient disappearance of the network model. The non-intrusive load decomposition based on DRANet uses the combined power data  $X$  as input samples and the power data  $Y$  of individual load devices as output labels to train and tune the network model. When tested on the test set, the decomposition results will be predicted based on the mapping relationship between  $X$  and  $Y$  during the training of the network. The model's feature enhancement module and information fusion module achieve optimal integration of load features extracted at different stages, thus improving the accuracy of the network model for load decomposition applications, especially improving the decomposition and identification of low-use appliances.

### 3 Simulation and analysis

#### 3.1 Data sets

In this paper, two datasets [UK-DALE public data set and WikiEnergy data set (Kelly and Knottenbelt, 2015)] are used to verify the algorithm. WikiEnergy data set is a research power data set released by Pecan Street, and it is the most abundant residential power energy database in the world to study power load decomposition. It contains power data collected by nearly 600 household users over a period of time, including single load and total household power consumption. The active power of all loads and residential buildings is obtained at the sampling frequency of 1/60 Hz. The collection of power data began in 2011, but it has not stopped. The database is still expanding, providing a good data support for the research of non-intrusive load decomposition. The UK-DALE public data mainly contains the information of single load and total household power consumption of five household

**TABLE 1** Load decomposition evaluation metrics comparison on WikiEnergy data sets.

Index	Method	Air condition	Fridge	Microwave	Washing machine	Dish washer
MAE	KNN	38.484	43.014	26.928	16.677	30.630
	DAE	36.964	39.520	17.015	12.081	25.107
	FHMM	40.212	22.898	12.787	12.787	19.782
	CNN s-p	39.635	13.760	13.155	11.959	11.624
	CNN s-s	61.129	38.413	9.973	18.497	19.084
	CD-LSTM	39.921	19.989	7.898	10.989	9.989
	D-ResNet	38.881	16.787	8.878	12.762	9.727
	DRAnet	<b>36.388</b>	<b>10.841</b>	<b>4.879</b>	<b>1.894</b>	<b>1.948</b>
SAE	KNN	0.0006	0.096	2.160	2.323	1.121
	DAE	<b>0.0001</b>	0.071	2.317	2.835	1.405
	FHMM	0.004	0.076	0.452	1.002	0.926
	CNN s-p	0.006	0.074	0.319	2.467	0.098
	CNN s-s	0.013	0.051	0.060	3.925	0.886
	CD-LSTM	0.065	0.045	0.172	1.190	0.158
	D-ResNet	0.011	0.021	0.182	1.022	0.152
	DRAnet	0.011	<b>0.019</b>	<b>0.133</b>	<b>0.08</b>	<b>0.084</b>

Bold indicates the best result.

users. The number of load devices in each household is up to 9, but the sampling period of each household is different. In the experimental, 70% of the sequences are used for training and 30% for testing.

In deep learning, artificially adding some noisy “dummy samples” combined with real data is beneficial to improve the robustness of the model and the generalization performance of the model. In this work, differential processing is performed on the raw power data. The load power signal is essentially a set of time-varying data, similar to a set of linear time-series information, and the total power samples are subjected to first-order differential processing, where the differential signal is a representation of the difference between two data. After the processing, the performance state of each non-zero-valued load device in the time dimension changes. Main purposes: 1) Each non-zero value power is changed to eliminate data fluctuation of load power signal and make the data tend to smoothness. 2) Combining raw data and differential data as the data input side enhances the data expression capability. The equation for differential processing is shown as follow Equation.

$$\Delta X_t = X_t - X_{t-1}, \quad (5)$$

where  $X_t$  represents the instantaneous total power data at the current moment  $t$ ,  $X_{t-1}$  represents the instantaneous total power data at the moment point  $t - 1$ ,  $\Delta X_t$  is the result of differential processing.

### 3.2 Analysis of experimental results for the WikiEnergy dataset

For the network model proposed in this work, experimental simulations and analyses are performed using several load devices,

namely, air conditioner, refrigerator, microwave oven, washing machine, and dishwasher in WikiEnergy data set. DRAnet conducts non-invasive load decomposition experiments on the corresponding load devices on the WikiEnergy dataset, and compares K-Nearest Neighbor (KNN) algorithm, factorial hidden Markov model (FHMM), Denoising AutoEncoder (DAE) algorithm, CNN Sequence to Sequence (CNN s-s) algorithm, and CNN Sequence to Point (CNN s-p) algorithm, composite deep long short-term memory network (CD-LSTM) (Xia et al., 2020), and deep dilated residual network (D-ResNet). In this work, the input size of the network is 100. For the sequence-to-sequence method, the output size is 100. For the sequence to point method, the output size is 1. The average absolute error (MAE) and the comprehensive absolute error (SAE) are used to evaluate the performance of the algorithm. **Table 1** and **Figure 8** shows the comparison between the load decomposition effects of the above algorithms and the real power data on the WikiEnergy dataset.

The experimental results show that the KNN algorithm has the worst load decomposition effect on the following loads, such as microwave oven, washing machine and dishwasher. Because these loads have the characteristics of low frequency use, the KNN algorithm cannot effectively identify and decompose the sudden change point of load power due to its own algorithm structure. All algorithms can effectively decompose the air conditioning load with periodic laws. From the decomposition results of refrigerators, it can be found that D-Resnet algorithm and DRAnet network model proposed in this paper can be decomposed better, mainly because they can accurately identify the peak area of load power. In fact, for load decomposition, the most important thing is to improve the decomposition ability at high power consumption. For the moment of very low power, although all methods are different, they have little

**TABLE 2** Load operating state index comparison on WikiEnergy data sets.

Index	Method	Air condition	Fridge	Microwave	Washing machine	Dish washer
Recall	KNN	0.998	0.996	0.759	0.290	0.561
	DAE	0.999	0.996	<b>1</b>	0.451	0.833
	FHMM	0.999	0.996	0.912	0.174	0.674
	CNN s-p	0.999	1	0.987	0.129	0.596
	CNN s-s	0.999	0.990	0.949	0.290	0.868
	CD-LSTM	0.999	0.994	0.957	0.453	0.856
	D-ResNet	1	0.997	0.947	0.675	0.887
	DRAnet	<b>1</b>	<b>1</b>	0.988	<b>0.880</b>	<b>0.982</b>
Precision	KNN	0.987	0.870	0.198	0.236	0.336
	DAE	0.987	0.853	0.050	0.229	0.281
	FHMM	0.991	0.966	0.061	0.132	0.421
	CNN s-p	0.995	0.996	0.050	0.047	0.414
	CNN s-s	0.939	0.847	0.033	0.428	0.391
	CD-LSTM	0.976	0.902	0.213	0.562	0.415
	D-ResNet	0.986	0.976	0.453	0.612	0.568
	DRAnet	<b>0.999</b>	<b>0.997</b>	<b>0.712</b>	<b>0.758</b>	<b>0.823</b>
Accuracy	KNN	0.991	0.889	0.967	0.993	0.978
	DAE	0.991	0.872	0.812	0.992	0.967
	FHMM	0.987	0.877	0.854	0.993	0.971
	CNN s-p	0.997	0.980	0.816	0.986	0.982
	CNN s-s	0.958	0.864	0.729	0.995	0.978
	CD-LSTM	0.976	0.978	0.821	0.995	0.968
	D-ResNet	0.975	0.981	0.942	0.987	0.971
	DRAnet	<b>0.999</b>	<b>0.986</b>	<b>0.979</b>	<b>0.996</b>	<b>0.996</b>

Bold indicates the best result.

impact on practical applications. Therefore, effective decomposition of load peaks is particularly important. The experiment shows that the DAE algorithm has certain advantages in identifying and decomposing regions with zero power consumption. For low frequency load equipment such as microwave ovens, washing machines and dishwashers, CNN s-s and CNN s-p convolutional neural networks can not accurately realize power decomposition. In terms of MAE and SAE, decomposition errors are relatively large. The main reason is that the number of layers of these two network models is small, and the load feature extraction is insufficient. D-ResNet’s performance is better than CNN s-s and CNN s-p, but it still cannot accurately realize the decomposition. That is because, although the residual structure makes the network deeper and improves the feature extraction ability, the feature extraction of electrical appliances with less frequency of use is still insufficient. Due to the structural advantages of the model, DRAnet uses dilated convolution to deepen the receptive field of convolution kernel, capture more time series information of fine load characteristics, and improve the decomposition effect. Compared with the existing decomposition algorithms, the DRAnet network model has better decomposition effect and better decomposition performance. In

particular, the load decomposition curve is closer to the real power consumption curve on the load of microwave oven, washing machine, dishwasher and other low-frequency use.

After the load is decomposed, the start and stop status of the electrical appliances can be distinguished by the threshold value. The threshold values of five kinds of appliances are: air conditioner 100 W, refrigerator 50 W, washing machine 20 W, microwave oven 200 W, dishwasher 100 W. **Table 2** shows the comparison of the evaluation indexes of load operation status after load decomposition by each algorithm. In Precision and Accuracy indexes, the DRAnet network model achieves the best decomposition performance in each appliance. In Recall metrics, good metrics performance was also achieved on Air Wither, Refrigerator, Washing Machine and Dishwasher. For CNN s-s, CNN s-p, DAE, CD-LSTM and D-ResNet, combining the decomposition results of microwave oven, washing machine and dishwasher in **Figure 7** with the load operation state metrics Recall and Precision analysis, the actual power consumption of these three loads is significantly less in the proportion of samples in the on state compared to the other two loads. Therefore, from the indexes, for these low-frequency use load devices, none of them can accurately identify the operating state of the load switch. The



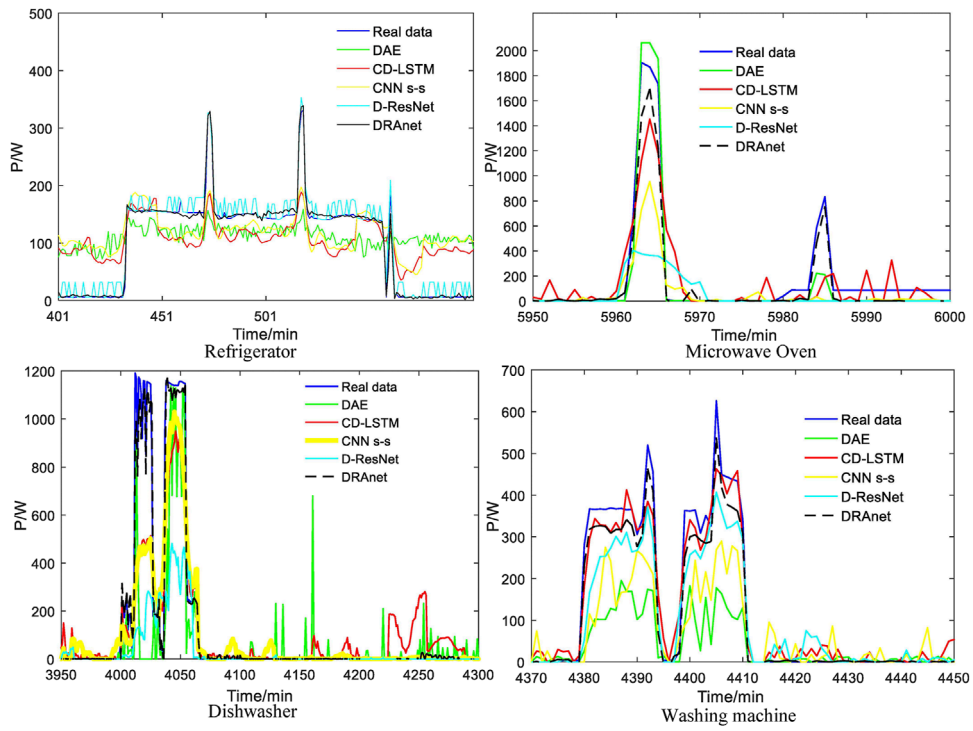


FIGURE 7 Partial view of load decomposition for WikiEnergy dataset.

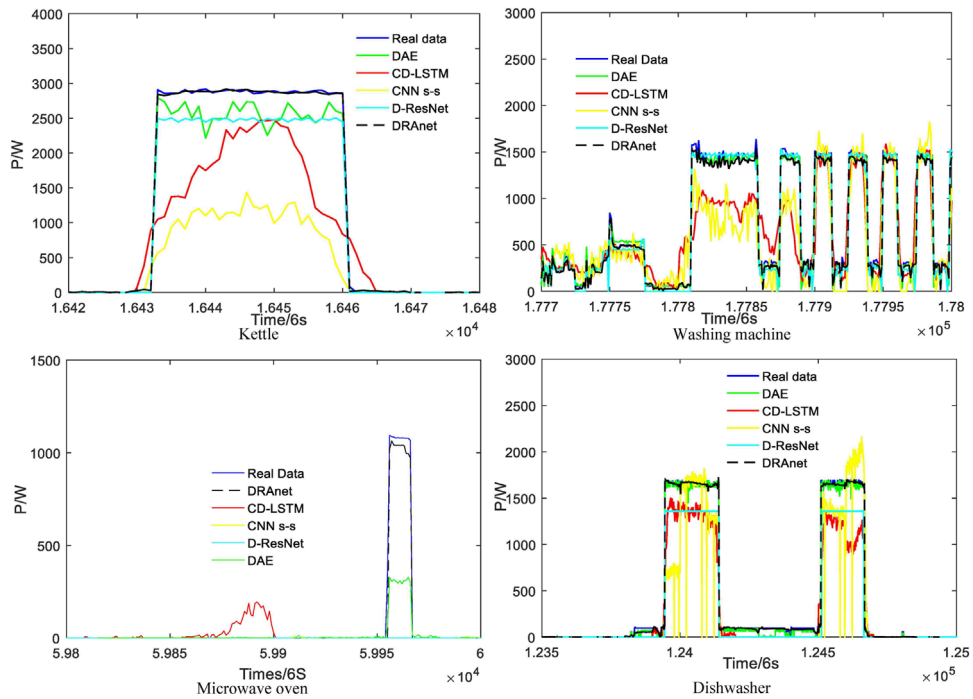


FIGURE 8 Partial view of load decomposition for UK-DALE dataset.

**TABLE 3** Load Decomposition Evaluation Metrics Comparison on UK-DALE data sets.

Index	Method	Kettle	Fridge	Microwave	Washing machine	Dish washer
MAE	KNN	1.413	2.407	1.078	4.032	3.274
	DAE	8.867	8.218	1.226	14.920	12.756
	FHMM	9.432	6.432	1.123	15.773	12.722
	CNN s-p	4.002	4.517	1.159	23.881	9.747
	CNN s-s	8.829	3.866	1.125	20.696	9.101
	CD-LSTM	3.139	3.762	1.021	17.622	7.653
	D-ResNet	3.028	2.862	0.972	14.526	6.525
	DRAnet	<b>0.703</b>	<b>2.133</b>	<b>0.811</b>	<b>3.618</b>	<b>2.403</b>
SAE	KNN	0.076	0.015	0.054	0.018	0.001
	DAE	0.377	0.021	0.748	0.006	0.340
	FHMM	0.004	0.076	0.452	1.002	0.926
	CNN s-p	0.242	0.024	0.845	0.302	0.154
	CNN s-s	0.522	0.032	0.880	0.315	0.213
	CD-LSTM	0.271	0.020	0.421	0.102	0.061
	D-ResNet	0.121	0.014	<b>0.226</b>	0.081	0.052
	DRAnet	<b>0.002</b>	<b>0.011</b>	0.237	<b>0.0007</b>	<b>0.020</b>

Bold indicates the best result.

best load identification is the network model proposed in this work, which is most capable of accurately predicting the operating state of such load switches. In general, the power of the refrigerator is relatively regular, that is, it will stay at a relatively high power position for a period of time after starting. Although there is a power peak during the period, this does not affect the judgment of the start and stop of the electric appliance.

### 3.3 Analysis of experimental results for the UK-DALE dataset

In order to verify the generalization performance of the network model in this paper, relevant comparative experiments were carried out in UK-DALE. In the electric power data, five typical loads of kettle, refrigerator, microwave oven, washing machine and dishwasher are selected for experiment and analysis. As shown in **Table 3**, the above five algorithms can achieve effective power decomposition for frequently used load equipment such as air conditioners and refrigerators. For loads such as washing machines, microwave ovens and dishwashers, KNN algorithm, FHMM algorithm, CNN s-s algorithm and CNN s-p algorithm are far less effective than CD-LSTM algorithm, D-ResNet algorithm and DRAnet network model. This is mainly due to the advantages of the proposed network structure, which can better detect the peak state of power consumption and the operating state change of load switch.

**Figure 8** shows the local load decomposition of UK-DALE data samples. Observe the decomposition curves of other methods. It is not well represented in the peak area of load power consumption,

and the curve has some burrs. But it performs well in the range of power close to zero. To sum up the four local decomposition renderings, the load decomposition result of DRAnet network model is the closest to the real power consumption compared with other algorithms, indicating that DRAnet decomposition performance is superior to other algorithms.

The decomposition performance of the algorithm is further evaluated by using the evaluation indexes of load start and stop operation states such as recall, accuracy and precision. As shown in **Table 4**, it is the comparison of load switch operation status indicators of five algorithms in UK-DALE data set. From the index point of view, DRAnet network model achieves the best numerical performance on these types of loads, and can accurately identify the operation status of load startup and shutdown. From the Recall index and Precision index, other algorithms perform well on kettles, dishwashers and refrigerators, and can also judge the start and stop status of electrical appliances, but they do not perform very well on other microwave ovens and washing machines. The accuracy of non-invasive load decomposition based on DRAnet is obviously superior to other methods.

On the two experimental results of WikiEnergy dataset and UK-DALE dataset, the dilated residual aggregated convolutional network based on its own structural advantages, the proposed dilated residual module enhances the network's ability to extract low-frequency load features, rescale the features using the feature enhancement module, filter the useless information, and reinforce the useful load features, thus to be have better decomposition effect and generalization performance than other methods.

TABLE 4 Load operating state index comparison on UK-DALE data sets.

Index	Method	Air condition	Fridge	Microwave	Washing machine	Dish washer
Recall	KNN	0.987	0.988	0.254	0.911	0.968
	DAE	0.985	0.944	0.345	0.921	0.938
	FHMM	0.992	0.937	0.362	0.931	0.941
	CNN s-p	0.993	0.923	0.425	0.838	0.928
	CNN s-s	0.969	0.990	0.463	0.857	0.904
	CD-LSTM	0.964	0.978	0.429	0.862	0.914
	D-ResNet	0.973	0.992	0.653	0.902	0.934
	DRAnet	<b>1</b>	<b>0.996</b>	<b>0.947</b>	<b>0.926</b>	<b>0.969</b>
Precision	KNN	0.998	0.974	0.933	0.617	0.799
	DAE	0.650	0.932	0.421	0.471	0.813
	FHMM	0.991	0.966	0.061	0.132	0.421
	CNN s-p	1	0.968	0.435	0.701	0.829
	CNN s-s	0.946	0.944	0.632	0.663	0.835
	CD-LSTM	0.964	0.964	0.643	0.673	0.801
	D-ResNet	0.973	0.976	0.732	0.701	0.823
	DRAnet	<b>1</b>	<b>0.996</b>	<b>0.989</b>	<b>0.771</b>	<b>0.850</b>
Accuracy	KNN	0.982	0.954	0.956	0.981	0.934
	DAE	0.997	0.955	0.949	0.967	0.976
	FHMM	0.972	0.954	0.954	0.943	0.965
	CNN s-p	0.987	0.981	0.967	0.936	0.967
	CNN s-s	0.988	<b>0.987</b>	0.966	0.955	0.974
	CD-LSTM	0.987	0.977	0.964	0.954	0.971
	D-ResNet	0.988	0.985	0.952	0.962	0.976
	DRAnet	<b>0.999</b>	0.986	<b>0.979</b>	<b>0.985</b>	<b>0.998</b>

Bold indicates the best result.

## 4 Summary

This work firstly introduces the overall structure of the dilated residual aggregated convolutional network, which mainly has the differential processing of data, the dilated residual module, the feature enhancement module, and the information fusion module. The differential data enhances the expressiveness of the data and improves the robustness of the network model, and then the null residual module, feature reinforcement module, and information fusion module are proposed for the characteristics of sparse features of low-usage load devices. In this work, the dilated residual aggregated convolutional network is trained and tested on the WikiEnergy dataset and UK-DALE dataset samples. Through experimental simulations and result analysis, the proposed method is significantly better than other methods. In terms of load sequence decomposition, the method proposed in this paper has significantly improved on MAE and SAE indicators compared with the existing deep learning decomposition method, and in terms of electrical start and stop judgment, the method in this paper is superior to the existing method in three indicators. The fundamental reason is

mainly that the dilated residual aggregated convolutional network model has a stronger extraction capability for higher order load features, and therefore has better decomposition results for low usage appliances. However, there are still many problems in non-intrusive load decomposition that need further study. 1) In order to improve the decomposition accuracy of load equipment, it is often to train the corresponding model for each load equipment. The process is complex and the time cost is high. Therefore, we can further study the deep neural network model with adaptive learning ability, such as the confrontation network model, which can be transplanted to the application of non-invasive load decomposition. 2) The decomposition of low power electrical appliances is more susceptible to noise interference and is not easy to decompose. The decomposition of low power loads needs further research. 3) The non-intrusive load decomposition algorithm based on deep learning in this paper has certain requirements for computing resources, and cannot be integrated with smart meters at present. In the future research work, we can consider further improvement of the algorithm, so that the algorithm can be directly used on embedded platforms and other hardware devices.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://paperswithcode.com/dataset/uk-dale>.

## Author contributions

FY: methodology, conceptualization, investigation; ZW: writing review and editing, methodology, validation; XZ: formal analysis, validation, resources; MX: supervision, funding acquisition, project administration.

## Funding

This work was supported by the Science and Technology Project of SGCC, named Research on the Construction Technology of Artificial Intelligence Application Verification Platform for Power Grid Dispatching in Typical Scenarios (SGSH0000DKJS2200274).

## References

- Ahmadi, H., and Marti, J. R. (2015). Load decomposition at smart meters level using eigenloads approach. *IEEE Trans. Power Syst.* 30, 3425–3436. doi:10.1109/tpwrs.2014.2388193
- Chang, H. H., Lin, L. S., Chen, N., and Lee, W. J. (2013). Particle-swarm-optimization-based nonintrusive demand monitoring and load identification in smart meters. *IEEE Trans. Industry Appl.* 49, 2229–2236. doi:10.1109/tia.2013.2258875
- Chen, B., Xia, M., Qian, M., and Huang, J. (2022). Manet: A multi-level aggregation network for semantic segmentation of high-resolution remote sensing images. *Int. J. Remote Sens.* 43, 5874–5894. doi:10.1080/01431161.2022.2073795
- Dash, S., and Sahoo, N. (2022). Electric energy disaggregation via non-intrusive load monitoring: A state-of-the-art systematic review. *Electr. Power Syst. Res.* 213, 108673. doi:10.1016/j.epr.2022.108673
- Fan, W., Liu, Q., Ahmadvpour, A., and Gholami Farkoush, S. (2021). Multi-objective non-intrusive load disaggregation based on appliances characteristics in smart homes. *Energy Rep.* 7, 4445–4459. doi:10.1016/j.ejegy.2021.07.033
- Gao, J., Weng, L., Xia, M., and Lin, H. (2022). MLNet: Multichannel feature fusion lozenge network for land segmentation. *J. Appl. Remote Sens.* 16, 1–19. doi:10.1117/1.jrs.16.016513
- Guo, Y., Xiong, X., Fu, Q., Xu, L., and Jing, S. (2021). Research on non-intrusive load disaggregation method based on multi-model combination. *Electr. Power Syst. Res.* 200, 107472. doi:10.1016/j.epr.2021.107472
- Hart, G. W. (1992). Nonintrusive appliance load monitoring. *Proc. IEEE* 80, 1870–1891. doi:10.1109/5.192069
- Hassan, T., Javed, F., and Arshad, N. (2014). An empirical investigation of v-i trajectory based load signatures for non-intrusive load monitoring. *IEEE Trans. Smart Grid* 5, 870–878. doi:10.1109/tsg.2013.2271282
- Himeur, Y., Alsalemi, A., Bensaali, F., and Amira, A. (2020a). Effective non-intrusive load monitoring of buildings based on a novel multi-descriptor fusion with dimensionality reduction. *Appl. Energy* 279, 115872. doi:10.1016/j.apenergy.2020.115872
- Himeur, Y., Alsalemi, A., Bensaali, F., and Amira, A. (2020b). An intelligent nonintrusive load monitoring scheme based on 2d phase encoding of power signals. *Int. J. Intelligent Syst.* 36, 72–93. doi:10.1002/int.22292
- Himeur, Y., Alsalemi, A., Bensaali, F., and Amira, A. (2020c). Robust event-based non-intrusive appliance recognition using multi-scale wavelet packet tree and ensemble bagging tree. *Appl. Energy* 267, 114877. doi:10.1016/j.apenergy.2020.114877
- Himeur, Y., Alsalemi, A., Bensaali, F., Amira, A., and Dimitrakopoulos, G. (2021a). “On the applicability of 2D local binary patterns for identifying electrical appliances in non-intrusive load monitoring,” in *Intelligent systems and applications*. Editor K. Arai (Berlin: Springer Nature), 188–205.
- Himeur, Y., Alsalemi, A., Bensaali, F., and Amira, A. (2021b). Smart non-intrusive appliance identification using a novel local power histogramming descriptor with an improved k-nearest neighbors classifier. *Sustain. Cities Soc.* 67, 102764. doi:10.1016/j.scs.2021.102764
- Himeur, Y., Alsalemi, A., Bensaali, F., and Amira, A. (2021c). “Appliance identification using a histogram post-processing of 2d local binary patterns for smart grid applications,” in International Conference on Pattern Recognition, Milan, Italy, 10–15 January 2021.
- Hu, K., Li, M., Xia, M., and Lin, H. (2022). Multi-scale feature aggregation network for water area segmentation. *Remote Sens.* 14, 206. doi:10.3390/rs14010206
- Inagaki, S., Egami, T., Suzuki, T., Nakamura, H., and Ito, K. (2011). Nonintrusive appliance load monitoring based on integer programming. *Electr. Eng. Jpn.* 174, 18–25. doi:10.1002/eej.21040
- Jia, Z., Yang, L., Zhang, Z., Liu, H., and Kong, F. (2021). Sequence to point learning based on bidirectional dilated residual network for non-intrusive load monitoring. *Int. J. Electr. Power & Energy Syst.* 129, 106837. doi:10.1016/j.ijepes.2021.106837
- Johnson, M. J., and Willsky, A. S. (2013). Bayesian nonparametric hidden semi-markov models. *J. Mach. Learn. Res.* 14, 673–701.
- Kelly, J., and Knottenbelt, W. (2015). “Neural nilm: Deep neural networks applied to energy disaggregation,” in 2nd ACM International Conference on Embedded Systems For Energy-Efficient Built Environments, Seoul, South Korea, November 4 - 5, 2015, 55–64.
- Kolter, J. Z., Batra, S., and Ng, A. Y. (2010). “Energy disaggregation via discriminative sparse coding,” in International Conference on Neural Information Processing Systems, British Columbia, Canada, 6–9 December 2010.
- Lin, Y., and Tsai, M. (2014). Development of an improved time–frequency analysis-based nonintrusive load monitor for load demand identification. *IEEE Trans. Instrum. Meas.* 63, 1470–1483. doi:10.1109/tim.2013.2289700
- Lin, Y. H., and Tsai, M. S. (2014). Non-intrusive load monitoring by novel neuro-fuzzy classification considering uncertainties. *IEEE Trans. Smart Grid* 5, 2376–2384. doi:10.1109/tsg.2014.2314738
- Lu, C., Xia, M., and Lin, H. (2022). Multi-scale strip pooling feature aggregation network for cloud and cloud shadow segmentation. *Neural Comput. Appl.* 34, 6149–6162. doi:10.1007/s00521-021-06802-0
- Luan, W., Yang, F., Zhao, B., and Liu, B. (2022). Industrial load disaggregation based on hidden markov models. *Electr. Power Syst. Res.* 210, 108086. doi:10.1016/j.epr.2022.108086
- Miao, S., Xia, M., Qian, M., Zhang, Y., Liu, J., and Lin, H. (2022). Cloud/shadow segmentation based on multi-level feature enhanced network for remote sensing imagery. *Int. J. Remote Sens.* 43, 5940–5960. doi:10.1080/01431161.2021.2014077

## Conflict of interest

Author ZW was employed by State Grid Shanghai Municipal Electric Power Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Park, S. W., Baker, L. B., and Franzone, P. D. (2019). Appliance identification algorithm for a non-intrusive home energy monitor using cogent confabulation. *IEEE Trans. Smart Grid* 10, 714–721. doi:10.1109/tsg.2017.2751465
- Paterakis, N. G., Erdinc, O., Bakirtzis, A. G., and Catalao, J. (2017). Optimal household appliances scheduling under day-ahead pricing and load-shaping demand response strategies. *IEEE Trans. Industrial Inf.* 11, 1509–1519. doi:10.1109/tii.2015.2438534
- Piga, D., Cominola, A., Giuliani, M., Castelletti, A., and Rizzoli, A. E. (2016). Sparse optimization for automated energy end use disaggregation. *IEEE Trans. Control Syst. Technol.* 24, 1044–1051. doi:10.1109/tcst.2015.2476777
- Qu, Y., Xia, M., and Zhang, Y. (2021). Strip pooling channel spatial attention network for the segmentation of cloud and cloud shadow. *Comput. Geosciences* 157, 104940. doi:10.1016/j.cageo.2021.104940
- Singh, S., and Majumdar, A. (2019). Analysis co-sparse coding for energy disaggregation. *IEEE Trans. Smart Grid* 10, 462–470. doi:10.1109/tsg.2017.2743763
- Singh, S., and Majumdar, A. (2018). Deep sparse coding for non-intrusive load monitoring. *IEEE Trans. Smart Grid* 9, 4669–4678. doi:10.1109/tsg.2017.2666220
- Song, L., Xia, M., Weng, L., Lin, H., Qian, M., and Chen, B. (2023). Axial cross attention meets cnn: Bibranch fusion network for change detection. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 16, 32–43. doi:10.1109/jstars.2022.3224081
- Tsai, M. S., and Lin, Y. H. (2012). Modern development of an adaptive non-intrusive appliance load monitoring system in electricity energy conservation. *Appl. Energy* 96, 55–73. doi:10.1016/j.apenergy.2011.11.027
- Wang, Z., Xia, M., Lu, M., Pan, L., and Liu, J. (2022). Parameter identification in power transmission systems based on graph convolution network. *IEEE Trans. Power Deliv.* 37, 3155–3163. doi:10.1109/TPWRD.2021.3124528
- Welikala, S., Dinesh, C., Ekanayake, M., Godaliyadda, R. I., and Ekanayake, J. (2019). Incorporating appliance usage patterns for non-intrusive load monitoring and load forecasting. *IEEE Trans. Smart Grid* 10, 448–461. doi:10.1109/tsg.2017.2743760
- Xia, D., Ba, S., and Ahmadpour, A. (2021). Non-intrusive load disaggregation of smart home appliances using the ippo algorithm and fhmm model. *Sustain. Cities Soc.* 67, 102731. doi:10.1016/j.scs.2021.102731
- Xia, M., Liu, W., Wang, K., Song, W., Chen, C., and Li, Y. (2020). Non-intrusive load disaggregation based on composite deep long short-term memory network. *Expert Syst. Appl.* 160, 113669. doi:10.1016/j.eswa.2020.113669
- Xia, M., Liu, W., Wang, K., Zhang, X., and Xu, Y. (2019). Non-intrusive load disaggregation based on deep dilated residual network. *Electr. Power Syst. Res.* 170, 277–285. doi:10.1016/j.epsr.2019.01.034
- Zhang, C., Zhong, M., Wang, Z., Goddard, N., and Sutton, C. (2016). Sequence-to-point learning with neural networks for nonintrusive load monitoring. Available at: <https://arxiv.org/abs/1612.09106>.
- Zhao, H., Yan, X., and Ren, H. (2019). Quantifying flexibility of residential electric vehicle charging loads using non-intrusive load extracting algorithm in demand response. *Sustain. Cities Soc.* 50, 101664. doi:10.1016/j.scs.2019.101664