International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2025): 8.036

An Industrial Non-Intrusive Load Disaggregation Method Based on Deep Learning Using Active and Reactive Power Features

Houkun Han

North China Electric Power University, School of Control and Computer Engineering, Changping District, Beijing, China

Email: 120232227072[at]ncepu.edu.cn

Abstract: ***Non-intrusive load monitoring provides a cost-effective way to obtain detailed equipment level energy consumption in buildings and industrial facilities. Existing studies mainly focus on residential loads and often rely only on active power features, which limits performance in industrial environments where reactive power is significant. This paper proposes an industrial load disaggregation method based on deep learning using both active and reactive power features. A hybrid architecture combining convolutional neural networks and long short-term memory networks is designed to extract spatial and temporal power signal features, and a channel attention mechanism is introduced to enhance important feature channels. Reactive power is incorporated as auxiliary information to improve the representation of industrial equipment operating patterns. Experiments conducted on the HIPE industrial dataset demonstrate that the proposed approach improves disaggregation accuracy compared with existing methods using only active power. The results confirm that reactive power information significantly enhances industrial load disaggregation performance.***

Keywords: Non-intrusive Load Monitoring, industrial load disaggregation, deep learning, load power feature

1. Introduction

The improvement in social electrification and informatization has led to continuously growing energy demand, urgently necessitating enhanced energy efficiency [1]. In this context, the industrial sector, due to its substantial scale of electricity consumption, has become a critical focal point for improving overall energy efficiency. Non-intrusive Load Monitoring (NILM) [2] can obtain detailed energy consumption information of individual appliances using only the aggregated load data measured at the electrical entry point. This fine-grained energy information can enhance energy management capabilities and improve energy utilization efficiency. Unlike intrusive load monitoring, NILM does not require installing sensing devices on each appliance individually, offering advantages such as low deployment cost and convenient maintenance.

Early non-intrusive load monitoring solutions relied on manually crafted appliance features, such as optimization methods [3], [4]. However, manual feature extraction for electrical loads is not only time consuming but also requires domain knowledge of the specific problem. In recent years, with the advancement of deep learning, various advanced algorithms have been proposed [5-6]. Kelly et al. [7] applied three deep neural network architectures to the load disaggregation task: the first was based on Long Short-Term Memory networks, the second was a denoising autoencoder, and the third was a regression network that predicted the start time, end time, and average active power of each appliance activation. Zhang et al. [8] proposed sequence-to-sequence (S2S) and sequence-to-point (S2P) methods. Among them, the S2P method, given an input window of aggregated active power signal, outputs only the active power value of the target appliance at the midpoint of the window, thereby reducing disaggregation error. Kaselim et al. [9] employed a bidirectional Long Short-Term Memory network (LSTM) to capture temporal features and accomplish the load disaggregation task.

Currently, NILM research predominantly concentrates on residential environments, while studies on industrial loads remain relatively underexplored. This research gap can be attributed to two primary factors: the difficulty in acquiring industrial load data, and the greater complexity of industrial equipment operation, which poses challenges for feature extraction [10]. Most residential appliances are either two-state devices (type-1) or multi-state devices (type-2). In contrast, industrial settings involve numerous continuously variable devices (type-3), which are inherently more difficult to characterize. Furthermore, industrial users follow unique production processes, resulting in load patterns with stronger temporal dependencies and higher noise levels compared to residential loads. Consequently, disaggregation methods developed for residential settings face inherent limitations when directly applied to industrial scenarios [11].

To address the challenges of industrial load disaggregation, this paper proposes a method based on deep learning utilizing both active and reactive power features. First, a hybrid model combining Convolutional Neural Networks and Long Short-Term Memory networks is constructed to extract power signal characteristics, incorporating a channel attention module to enhance critical features. Second, reactive power features are introduced as supplementary information to enable more comprehensive capture of equipment operational patterns. Unlike current deep learning solutions that typically employ only active power, this approach leverages reactive power as an additional feature, given that many industrial loads are not purely resistive and generate substantial reactive power during operation. Finally, experiments conducted on the industrial electricity dataset HIPE validate the effectiveness of the proposed method.

2. Non-intrusive Load Disaggregation Concept

For a given industrial power user, the total load active power sequence X={ x (t)}T t=1 collected at the user's incoming line at time t can be expressed as:

$$x(t)=\sum_{i=1}^{N} y_i(t)+e(t) \tag{1}$$

where x (t) represents the total active power at time t, y_i (t) denotes the active power of the i-th appliance at time t, e (t) is the measurement error, and N is the total number of devices.

When performing the load disaggregation task, the output signal of the model refers to the active power data of the target appliance. For the input signal, the NILM field typically only employs the active power data collected at the bus side. However, in addition to active power input, this paper introduces reactive power information as an additional input.

In the proposed method, let the aggregated active power input to the disaggregation model be $P_{t:t+T-1}$ and the aggregated reactive power be $Q_{t:t+T-1}$. The disaggregation model maps these two types of inputs to the active power $p_{i,t:t+T-1}$ of the i-th appliance, with the mapping relationship expressed as:

$$p_{i,t:t+T-1}=f\left(P_{t:t+T-1},Q_{t:t+T-1}\right) \tag{2}$$

where T is the total length of the input time series window, t is the starting time of the input window, and f is the mapping function learned by the model.

3. Proposed Method

3.1 Overall network architecture

The overall architecture of the proposed industrial non-intrusive load disaggregation network is illustrated in **Figure 1**. The network takes the aggregated active power sequence P and reactive power sequence Q in industrial scenarios as inputs, and achieves accurate disaggregation of the target appliance's active power sequence through the collaborative design of feature extraction layer and regression layer.

First, the input aggregated power sequences P and Q are concatenated to integrate the active and reactive signals into a unified two-dimensional feature map, which is then fed into the subsequent network. Then, deep features are extracted through N serial feature extraction layers. Each feature extraction layer consists of a convolutional layer and a channel attention Squeeze-and-Excitation Network (SENet) module. The convolutional layer is responsible for primary feature extraction, while the SENet attention mechanism explicitly models interdependencies between channels to automatically enhance feature channels critical for the load disaggregation task and suppress redundant information interference. The proposed network stacks five feature extraction layers (i.e., N=5). From the first to the fifth layer, the kernel sizes are 2×10, 2×8, 2×6, 2×5, and 2×5, respectively; the numbers of kernels are 30, 30, 40, 50, and 50, respectively. Each convolutional layer is followed by a ReLU activation function before being fed into the SENet.

After the output of the final feature extraction layer, the feature map is reshaped to match the input format of the LSTM. The final load disaggregation task is completed through a regression layer, which consists of an LSTM and a fully connected layer. The LSTM performs deep temporal modeling of features at each time step to capture the temporal dependencies of load power variations, and the fully connected layer ultimately maps the output to the entire active power sequence of the target appliance within the time window.

The loss function adopts the Mean Squared Error (MSE), which is calculated as:

$$Loss=\frac{1}{T}\sum_{t=1}^{T}\left(\hat{p}_i(t)-p_i(t)\right)^2 \tag{3}$$

where T is the number of time points, $\hat{p}_i(t)$ and $p_i(t)$ are the predicted active power value of appliance and the ground truth active power value of appliance i at time t.

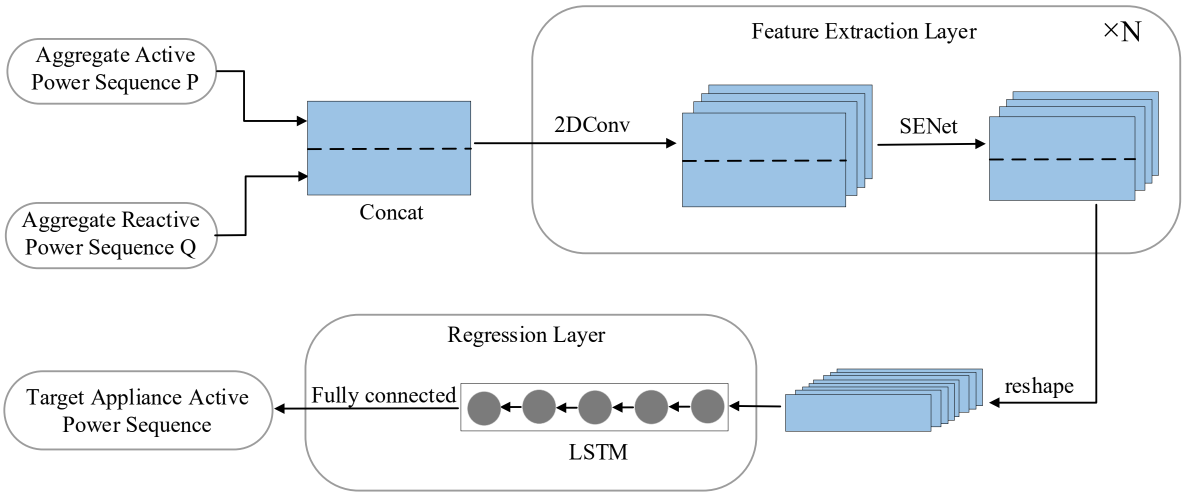

Figure 1: Overall network architecture

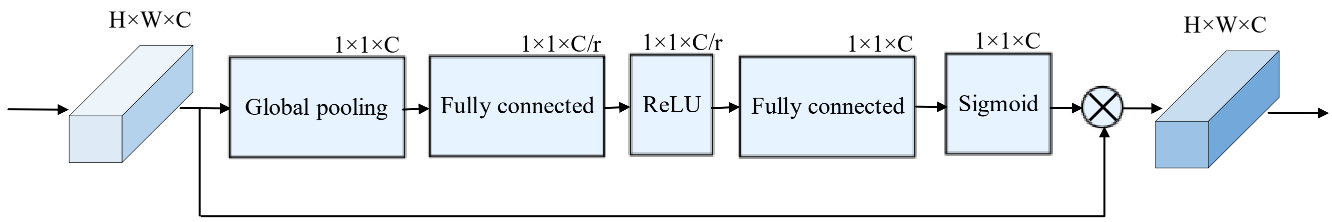

Figure 2: SENet architecture

3.2 SENet channel attention

The output of the convolutional neural network generates a large number of feature channels. In the feature extraction layers of the proposed network, a SENet (Squeeze-and-Excitation Network) [12] channel attention module is embedded after each convolutional layer. SENet is a lightweight channel attention mechanism that models the relationships between feature map channels through squeeze and excitation operations, thereby enhancing feature channels critical to the task. The SENet structure adopted in this paper is illustrated in **Figure 2**.

Squeeze operation: Global average pooling is performed on the input feature map, transforming the feature map of shape H×W×C into a shape of 1×1×C, where C is the number of channels, and H and W are the height and width of the feature map, respectively. This operation achieves global aggregation of channel information.

$$z = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i, j) \tag{4}$$

where z is the squeeze output, u_c is the input matrix, and i, j denotes its row and column indices, respectively.

Excitation operation**:** The excitation operation consists of two fully connected layers, which enhance the feature aggregation capability through ReLU and Sigmoid activation functions, respectively. This design enables the model to automatically learn the importance of each channel.

$$s = \sigma\left(W_2 \delta\left(W_1 z\right)\right) \tag{5}$$

where s is the excitation output, σ denotes the Sigmoid activation function, δ represents the ReLU activation function, and W_1 and W_2 are the weights of the two fully connected layers.

3.3 LSTM

Leveraging the advantages of LSTM in handling time series problems, this paper feeds the output of the feature extraction layer into the LSTM network within the regression layer to learn the relationship between sequential features and load active power. LSTM is composed of multiple cells, each containing a forget gate, an input gate, and an output gate. Its structure is illustrated in **Figure 3**.

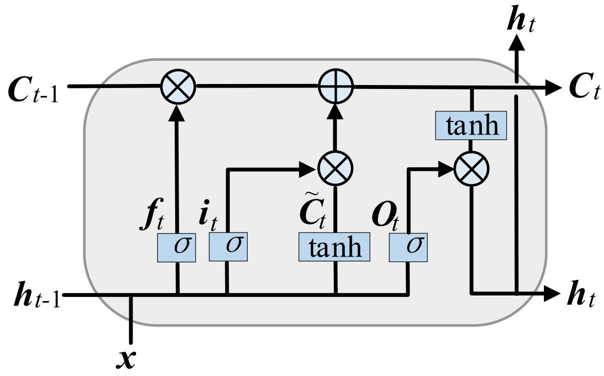

Figure 2: LSTM architecture

The specific calculation procedure of LSTM is as follows: The forget gate controls the retention ratio of the previous cell state:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t\right] + b_f\right) \tag{6}$$

The input gate controls the update ratio of the current input information:

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t\right] + b_i\right) \tag{7}$$

$$\tilde{C}_t = \tanh\left(W_c \cdot \left[h_{t-1}, x_t\right] + b_c\right) \tag{8}$$

The cell state is updated by combining the forget gate and input gate:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{9}$$

The output gate controls the output of the cell state:

$$O_t = \sigma\left(W_O \cdot \left[h_{t-1}, x_t\right] + b_O\right) \tag{10}$$

$$h_t = O_t \odot \tanh\left(C_t\right) \tag{11}$$

where C_t is the cell state at time step t; h_t is the cell output at time step t; x_t is the cell input at time step t; f_t is the output of the forget gate at time step t; i_t and $\tilde{C}_t$ jointly represent the output of the input gate at time step t; O_t is the output of the output gate at time step t; σ denotes the sigmoid activation function, and W and b are learnable weights and biases.

4. Experimental Results and Analysis

4.1 Dataset Introduction and Preprocessing

The HIPE dataset [13] contains equipment energy consumption data from an electronics factory in Germany over a three-month period from October 1, 2017, to January 1, 2018. The sampling frequency is 1/5 Hz, and each data point includes various electrical parameters such as active power, reactive power, and voltage. This chapter focuses on disaggregating the following four types of equipment from the dataset: Screen Printer, Large Washing Machine, High-Temperature Oven, and Soldering Oven. As shown in **Table 1**, these appliances encompass type-1, type-2, and type-3 categories.

Table 1: Appliance information

Appliance	Type	Abbreviation
ScreenPrinter	type-1	SP
WashingMachine	type-2	WM
HighTemperatureOven	type-3	HTO
SolderingOven	type-3	SO

For data processing, the dataset was first resampled at a frequency of 1/5 Hz. Missing values were imputed using forward filling, and data segments with consecutive missing values exceeding 30 sampling points were removed.

4.2 Evaluation Metrics

This paper adopts Mean Absolute Error (MAE) and Normalized Disaggregation Error (NDE) as evaluation metrics to assess model performance.

MAE measures the absolute difference between predicted and ground truth values:

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|\hat{p}_i(t) - p_i(t)\right| \tag{12}$$

NDE is more sensitive to larger errors in predictions and intuitively reflects the relative deviation between predicted and ground truth values:

$$NDE = \frac{\sum_{t=1}^{T}\left(\hat{p}_i(t) - p_i(t)\right)^2}{\sum_{t=1}^{T} p_i(t)^2} \tag{13}$$

where T is the number of time points, $\hat{p}_i(t)$ and $p_i(t)$ are the predicted active power value of appliance and the ground truth active power value of appliance i at time t.

4.3 Training Settings

The processed data was split into training, validation, and test sets in a ratio of 5:2:3. The sliding window length was set to 599. Additionally, Z-Score normalization was applied to the data.

During training, the batch size was set to 100, the maximum number of epochs was set to 60, and an early stopping mechanism was introduced to prevent overfitting and reduce training time. The Adam optimizer was employed with an initial learning rate of 0.001.

4.4 Experimental Results

This paper compares the performance of the proposed method with commonly used NILM methods S2P and S2S. Among these, S2P and S2S use only aggregated active power data as input, while the proposed method utilizes both aggregated active power and reactive power data as input. Additionally, an ablation study was conducted using a variant of the proposed model that employs only active power as input (denoted as Proposed (Only P)). In this variant, the reactive power input and feature concatenation module were removed, and the shapes of network layers were adjusted accordingly to assess the value of introducing reactive power. The complete proposed method is named Proposed (P and Q). **Table 2** presents the disaggregation performance comparison of different methods.

Table 2: Experimental results of each method

Appliance	Method	MAE/kW	NDE
SP	S2P	0.047	0.655
	S2S	0.052	0.681
	Proposed (Only P)	0.046	0.637
	Proposed (P and Q)	0.040	0.601
WM	S2P	0.052	0.774
	S2S	0.044	0.614
	Proposed (Only P)	0.041	0.562
	Proposed (P and Q)	0.039	0.573
HTO	S2P	0.185	0.632
	S2S	0.182	0.643
	Proposed (Only P)	0.168	0.612
	Proposed (P and Q)	0.155	0.571
SO	S2P	0.067	0.421
	S2S	0.069	0.398
	Proposed (Only P)	0.063	0.380
	Proposed (P and Q)	0.057	0.341

From the experimental results in **Table 2**, it can be observed that the proposed complete method (Proposed (P and Q)) achieves the best overall disaggregation performance. In terms of the MAE metric, Proposed (P and Q) shows significant improvement compared to existing methods S2P and S2S. Taking HTO as an example, the MAE values of S2P and S2S are 0.185kW and 0.182kW, respectively, while Proposed (P and Q) reduces it to 0.155kW. Regarding the NDE metric, Proposed (P and Q) achieves the best results on three types of appliances: SP, HTO, and SO. Particularly on the SO appliance, the NDE value reaches 0.341, significantly lower than S2S's 0.398, indicating that the proposed method yields more stable prediction results and demonstrating the effectiveness of the proposed approach.

By comparing the performance of Proposed (only P) and Proposed (P and Q), the contribution of reactive power features to load disaggregation can be clearly evaluated. On the SP appliance, the introduction of reactive power reduces MAE from 0.046kW to 0.040kW and NDE from 0.637 to 0.601. On the HTO appliance, MAE decreases from 0.168kW to 0.155kW and NDE from 0.612 to 0.571. On the SO appliance, MAE drops from 0.063kW to 0.057kW and

NDE from 0.380 to 0.341. On the WM appliance, although the NDE metric of Proposed (P and Q) (0.573) is slightly higher than that of Proposed (only P) (0.562), the MAE metric improves from 0.041kW to 0.039kW. These results indicate that the introduction of reactive power features can effectively enhance model disaggregation accuracy, with particularly significant improvement in the NDE metric. This is because industrial equipment operation involves continuous reactive power exchange, and this information provides auxiliary features beyond active power for the model, contributing to a more comprehensive learning of equipment operation patterns.

From the perspective of appliance types, the proposed method shows the most significant performance improvement on type-3 appliances (HTO, SO). These two types belong to continuously variable appliances with higher disaggregation difficulty, and the introduction of reactive power provides rich auxiliary information for the model. Based on the comprehensive analysis above, it can be concluded that reactive power features play a significant auxiliary role in industrial load disaggregation, and the synergistic effect of both features outperforms single active power input, with the most notable improvement observed in type-3 appliances. This demonstrates that the proposed method effectively enhances industrial load disaggregation performance.

5. Conclusion

This study proposes a deep learning based industrial load disaggregation method that integrates active and reactive power features. The model combines convolutional layers, channel attention through SENet, and LSTM based temporal modeling to capture complex operational patterns of industrial equipment. Experimental evaluation on the HIPE dataset demonstrates that incorporating reactive power improves disaggregation accuracy compared with conventional methods that rely only on active power. The results confirm the importance of reactive power as an auxiliary feature for industrial NILM tasks. Future work may explore additional electrical features and multi task learning strategies to further improve robustness across diverse industrial environments.

References

[1] Tanoni G, Principi E, Squartini S. Non-intrusive load monitoring in industrial settings: A systematic review[J]. Renewable and Sustainable Energy Reviews, 2024, 202: 114703.
[2] Hart G W. Nonintrusive appliance load monitoring[J]. Proceedings of the IEEE, 1992, 80 (12): 1870-1891.
[3] Wu X, Jiao D, You L. Nonintrusive on-site load-monitoring method with self-adaption[J]. International Journal of Electrical Power & Energy Systems, 2020, 119: 105934.
[4] Lin Y H. Trainingless multi-objective evolutionary computing-based nonintrusive load monitoring: Part of smart-home energy management for demand-side management[J]. Journal of Building Engineering, 2021, 33: 101601.
[5] Ding D, Li J, Zhang K, et al. Non-intrusive load monitoring method with inception structured CNN[J]. Applied Intelligence, 2022, 52 (6): 6227-6244.
[6] Liu X, Jiang T, Luo H. GI-Transformer: a transformer model with feature interaction for Non-Intrusive load monitoring[J]. IEEE Transactions on Instrumentation and Measurement, 2025.
[7] Kelly J, Knottenbelt W. Neural nilm: Deep neural networks applied to energy disaggregation[C]//Proceedings of the 2nd ACM international conference on embedded systems for energy-efficient built environments. 2015: 55-64.
[8] Zhang C, Zhong M, Wang Z, et al. Sequence-to-point learning with neural networks for non-intrusive load monitoring[C]//Proceedings of the AAAI conference on artificial intelligence. 2018, 32 (1).
[9] Kaselimi M, Doulamis N, Voulodimos A, et al. Context aware energy disaggregation using adaptive bidirectional LSTM models[J]. IEEE Transactions on Smart Grid, 2020, 11 (4): 3054-3067.
[10] Zhang Y, Chen J, Li C, et al. Non-Intrusive Industrial Load Monitoring Considering Load Power Characteristics and Timing Correlation[J]. IEEE Transactions on Instrumentation and Measurement, 2025.
[11] Kalinke F, Bielski P, Singh S, et al. An evaluation of nilm approaches on industrial energy-consumption data[C]//Proceedings of the twelfth ACM international conference on future energy systems. 2021: 239-243.
[12] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
[13] Bischof S, Trittenbach H, Vollmer M, et al. HIPE: An energy-status-data set from industrial production[C]//Proceedings of the ninth international conference on future energy systems. 2018: 599-603.

Paper ID: SE26303201043 DOI: https://dx.doi.org/10.70729/SE26303201043