Evaluation of Operating State for Smart Electricity Meters Based on Transformer–Encoder–BiLSTM

Zhongyong Zhao, *Member, IEEE*, Yu Chen, Jiangnan Liu, Yingying Cheng, Chao Tang, *Member, IEEE*, and Chenguo Yao, *Member, IEEE*

***Abstract*—The reliable operating state of smart electricity meters is significant in industrial applications. Faulty meters or meters in a poor measurement state will seriously impact both customers and stakeholders. However, the maintenance personnel now still use the manually periodic sampling inspection from the batch of smart electricity meters to evaluate the state of the entire batch, which has limitations of blindness, poor real time, inadequate, and insufficient examination. With the population of smart meters, it is feasible to evaluate the health condition of these devices with big data and artificial intelligence technology. One significant contribution of this article is first proposing an operating state evaluation method of smart electricity meters based on the transformer–encoder and bidirectional long-term and short-term memory. The evaluation indicators and preprocess of meters' data are carefully selected. A deep neural network is constructed and trained, the experimental verification is carried out, and the performance of the proposed method is compared with that of other traditional methods. The results show that the average classification accuracy of the proposed neural network model is 99.5%. Besides, compared with conventional machine learning and deep learning models, the proposed model is suitable for the operation state evaluation of smart electricity meters. From the experimental result, the potential benefit of the proposed method is that it could improve the accuracy and robustness of state evaluation.[1]**

***Index Terms*—Bidirectional long-term and short-term memory (BiLSTM), deep learning, neural network, smart electricity meters, state evaluation, transformer–encoder.**

Manuscript received 26 October 2021; revised 13 January 2022 and 11 April 2022; accepted 25 April 2022. Date of publication 3 May 2022; date of current version 3 March 2023. This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant SWU-KT22027, in part by the National Natural Science Foundation of China under Grant 51807166, in part by the Venture and Innovation Support Program for Chongqing Overseas Returnees under Grant cx2019123, and in part by the Natural Science Foundation of Chongqing under Grant cstc2019jcyj-msxmX0236. Paper no. TII-21-4692. *(Corresponding author: Zhongyong Zhao.)*

Zhongyong Zhao, Yu Chen, Jiangnan Liu, and Chao Tang are with the College of Engineering and Technology, Southwest University, Chongqing 400716, China (e-mail: zhaozy1988@swu.edu.cn; cy1034429543@email.swu.edu.cn; xjy7641@sina.com; tangchao_1981@163.com).

Yingying Cheng is with State Grid Chongqing Electric Power Company Marketing Service Center (Metrology Center), Chongqing 401120, China (e-mail: cyy_99@sina.com).

Chenguo Yao is with the State Key Laboratory of Power Transmission Equipment and System Security and New Technology, Chongqing University, Chongqing 400044, China (e-mail: yaochenguo@cqu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2022.3172182.

Digital Object Identifier 10.1109/TII.2022.3172182

I. Introduction

WITH the popularization of smart electricity meters by power grid companies all over the nation, smart electricity meters have become an indispensable part of the power grid [1]–[3]. The reliable operating state of smart electricity meters is significant for both customers and stakeholders. It requires that the electric energy be accurately measured by the smart electricity meters, regardless of various smart electricity meters. Thus, the state evaluation and fault diagnosis of smart electricity meters become necessary and a hotspot [4]. The state evaluation of the smart electricity meter is realized by advanced state monitoring means and reliable evaluation algorithms. The degree of operating state and development trend of the failure can be estimated, which can guide inspection and maintenance for staff before the performance of the smart meter drops to a certain extent.

Currently, the main focus of researchers on smart electricity meters is load forecasting [5], [6], analysis of user behavior [7]–[12], and detection of electricity theft [13]. However, there are few studies regarding the operating state evaluation of smart meters [14]–[16]. Nowadays, periodic inspection is often used to evaluate the states of the meters in engineering. Sample testing is used in periodic inspection. The results of the periodic inspection can be used to estimate the operating conditions of the entire batch of smart meters, about thousands to tens of thousands. Thus, sampling results provide the guidance that the whole bunch of meters either continue operating or are withdrawn from service. Although this routine method can achieve a specific effect, there are always some problems in the periodic inspection, such as blindness, poor real time, inadequate, and insufficient examination. Some smart electricity meters with good performance may have been forced to be out of service when they reach

[1]Code [Online]. Available: https://github.com/cy1034429432/Transfomer-encoder-bilstm-acc-99.5-.git

the deadline for rotation or due to the unqualified samples in the same batch, resulting in many wastes of resources. Besides, the deficiencies are hardly found for some other meters with certain defects because the inspection time is not reached, resulting in metering problems and economic losses [4].

Therefore, new thoughts and research have emerged in recent years to overcome the shortcomings of periodic inspection of smart electricity meters [4], [17], [18]. There are four methods proposed for evaluating the operating state of smart electricity meters as follows:

1) Establish a reasonable and empirical mathematical model to assess the reliability of smart electricity meters.
2) Install attached hardware to collect the information and monitor the status of the smart electricity meters in real time.
3) Build a comprehensive evaluation system according to the on-site environment and various quality indicators.
4) Develop data-driven evaluation methods for the operating state of smart electricity meters.

The first method pays attention to the structure of the smart electricity meter itself compared with other methods. Zhang *et al.* [19] established a failure rate model by component stress method, which can simulate the failure or abnormality of smart electricity meters based on hardware structure and function design. Based on the component manual, it calculates the meantime to failure to characterize smart electricity meters' life estimation. In Xu *et al.* [20], a multidegradation model of smart electricity meters based on Vine Copulas is proposed. It addresses the problem of lack of multirelevance in high-dimensional indicators and gives the joint distribution of reliability of smart electricity meters. However, the mathematical models usually use static parameters, while some information of smart meters is dynamic time-series data. The model cannot adapt to the changes of physical models of smart meters. Besides, the mathematical derivation is complex, and the simplification will lead to some errors. Moreover, this method needs to reconstruct the different smart electricity meters model. Thus, it has poor generalization ability.

In the second method, building an online and real-time attached hardware to construct the detection system is a useful way to maintain the stability of the meters [1]–[3]. However, this method requires a lot of extra material resources; the technique is still in the stage of development [4], [21]. Wei *et al.* [22] presented the integrated sensor environment and remote monitoring system for real-time monitoring of smart electricity meters. Feng *et al.* [23] constructed a multidimensional detection system for evaluating states of smart electricity meters based on the decision tree group. This system can send an early warning signal when the smart electricity meter is abnormal, which could assist the staff in replacing the abnormal meters. However, this method requires additional measuring equipment for smart electricity meters to receive information. It is almost impossible to install attached devices to every smart electricity meter, and the method is inefficient and uneconomic. Therefore, it is the most economical and feasible scheme to use the own sampling data from the smart meters instead of the recording data from attached hardware.

The third method is simpler than the other three methods. Usually, engineers combine this method with manual evaluation to ensure that the evaluation method has certain objectivity. Cheng *et al.* [17] proposed a comprehensive evaluation method based on entropy weight theory and grey correlation degree, which can evaluate the state of meters in four stages of the life cycle of meters: production supervision evaluation, predelivery quality evaluation, postarrival quality evaluation, and on-site operation evaluation. Ying *et al.* [18] proposed a fuzzy analytic hierarchy process and an adaptive weighting method to establish the state evaluation model of the smart electricity meter. The state can be evaluated based on the basic information of the smart electricity meters, the operation monitoring data, and the field measured values. Although the third method has been widely used in the power system, human factors affect the evaluating process, and further accuracy should be improved. Therefore, power companies are eager to improve and update this technique.

In the fourth method, with the progress of computing power and advanced algorithms, many researchers apply data-driven strategies to the analysis, diagnosis, decision-making, and management of smart electricity meters, based on machine learning. Jiao *et al.* [24] established a state evaluation model driven by validation data to obtain the error state of the smart electricity meters. Yip *et al.* [25] introduced two algorithms based on linear regression to study the electricity consumption behavior of consumers and detect the defective smart electricity meters. Helong *et al.* [26] and Feng *et al.* [27] use traditional machine learning methods, support vector machine, regression, decision tree, naive Bayes, etc. to perform the fault prediction and abnormal diagnosis of smart electricity meters. However, [26], [27] classify the smart meters based on the indicators proposed in the third method. Still, the human factors will affect the results.

Based on the above four known operating state evaluation methods of smart electricity meters, there are still the following problems in this field:

1) The method of evaluating the operating state of smart electricity meters is too complex, and most of them are offline. Although a few scholars use a relatively simple machine learning method for operating state evaluation, the indicators used in this method are too subjective to focus on the data of smart electricity meters themselves.
2) There is no unified standard for the fault detection of smart electricity meters. At present, it is still manual to estimate smart electricity meters, and there is no intelligent method.
3) Most of the current operating states evaluation models are of low accuracy and have not been widely used in power system online monitoring.

In summary, to make the evaluation methods more convenient and accurate [16], it is necessary to utilize the advantages of big data of electricity information and use deep learning to improve the automation and intelligence level of operating state evaluation [5]. For the first time, this article proposes transformer–encoder–bidirectional long-term and short-term memory (BiLSTM) and applies it to process extensive sampling data of smart electricity meters to evaluate their operating state. This article makes the following main contributions:

1) Based on the smart meter's data, this article proposes a new deep learning model, named transformer–encoder–BiLSTM, for the operating state evaluation of smart electricity meters. The input of this model focuses on the state data of the smart electricity meters themselves, which is more objective than other indicators.
2) This article proposes the operating state evaluation method of smart electricity meters based on the proposed model. The evaluation procedure governed by machines can replace manual inspection methods. Compared with the manual ways, this method can be intelligent, accurate, automatic, and fast, improving the reliability of electric energy metering. The unqualified and faulty smart electricity meters for stakeholders or power companies can be replaced by new smart meters based on the proposed method. It realizes the transition from "periodical verification-based replacement" to "state evaluation-based replacement" of smart electricity meters, saving time, human resources, and material costs. For society, this scheme will reduce the environmental pollution from electronic products.
3) The proposed model has a strong anti-interference ability to measurement error and noise, strong generalization ability, and higher accuracy than other methods. And then, the trained model can be integrated into existing online systems to realize online detection of smart electricity meters.

The rest of this article is organized as follows. Section II first introduces the transformer–encoder, long short-term memory (LSTM), and BiLSTM theory. The experimental setting, including selecting evaluation indicators, data preprocessing, and network's hyperparameter settings, is proposed in Section III. Section IV analyses the experimental results, comparing the effect of transformer–encoder–BiLSTM and other traditional methods. The suggestion, future analysis, and conclusion are summarized in Sections V and VI, respectively.

II. Theory of Neural Network for Deep Learning

A. Theory of Transformer–Encoder

In Fig. 1, it is a classic transformer–encoder structure [28]. It consists of a multihead attention layer, a feedforward neural network layer, and two ADD and norms layers. The multihead attention layer can aggregate the sequence data to extract feature information, but its weight can also reflect the impact of indicators on the results. Compared with the traditional recurrent neural network (RNN) structure, it has a faster operation speed and stronger feature extraction ability. The ADD structure is derived from Resnet, which can solve the gradient dispersion of the deep network. Norm refers to layer norm, which can speed up training. In recent years, this structure [29], [30] has greatly contributed to applying various fields.

B. Theory of Bidirectional Long Short-Term Memory

Many smart electricity meter data are time-series data; thus, the LSTM processes these sequence parameters. The overall

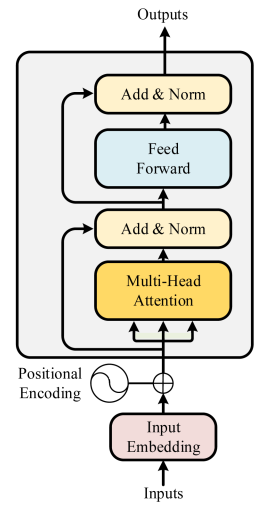

Fig. 1. Structure diagram of transformer–encoder.

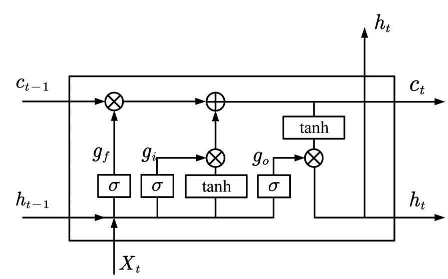

Fig. 2. Structure of LSTM neuron.

structure of LSTM is similar to that of RNN, but the structure of the repetitive modules within its neurons is much more complex, as shown in Fig. 2. It conveys both forward information and current information. LSTM introduces a new internal state to transfer linear cyclic information, outputs information to the hidden state, and elects to retain or forget information through gated recurrent units. The input gate controls how much input information needs to be kept at the current moment. The forget gate controls how much information needs to be discarded at the last moment, and the output gate controls how much information needs to be output to the hidden state h_t at the current moment. The updated formula for each time state of LSTM is as follows:

$$h_t = g_o^{(b)} f_h(C_t) \tag{1}$$

$$C_t = g_f^{(t)} C_{t-1} + g_i^{(t)} f_s(wh_{t-1} + uX_t + b) \tag{2}$$

$$softsign(x) = \frac{x}{1+|x|} \tag{3}$$

$$g_i^{(t)} = softsign(w_i h_{t-1} + u_i X_t + b_i) \tag{4}$$

$$g_f^{(t)} = softsign(w_f h_{t-1} + u_f X_t + b_f) \tag{5}$$

$$g_o^{(t)} = softsign\left(w_o h_{t-1} + u_o X_t + b_o\right) \tag{6}$$

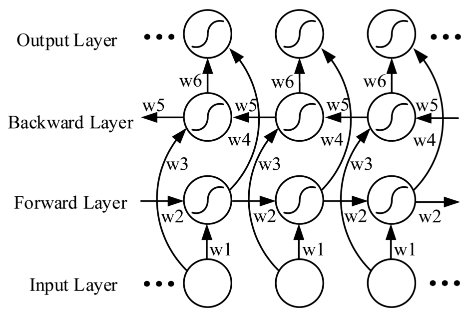

Fig. 3. Structure diagram of BiLSTM network.

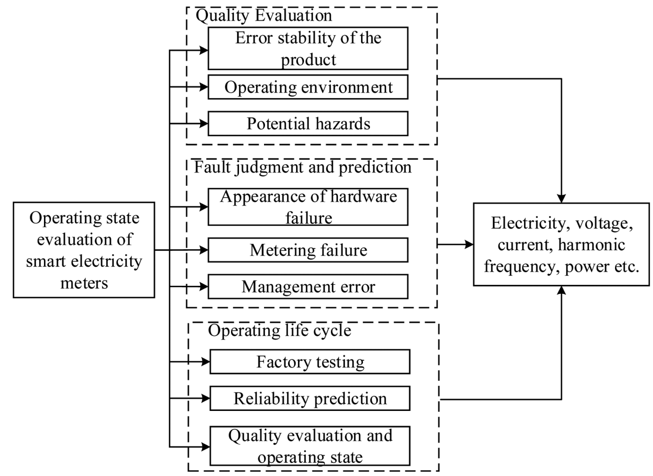

Fig. 4. Indicators that influence the state evaluation of smart electricity meters [21].

where f_h and f_s are activation functions of system state and internal state, it is usually hyperbolic tangent function tan*h*, *b* is offset constant, subscripts "*i*," "*f*," and "*o*" denote input gate, forget gate, and output gate, respectively, and *g* is a control gate unit updated with time step, which is essentially a feedforward neural network with *softsign* function as activation function. The *softsign* function has flatter curves and slower descent derivatives than the tan*h* and *sigmoid* functions, indicating that it can learn more efficiently and solve gradient disappearance better than the tan*h* and *sigmoid* functions.

To increase the impact of the historical sequence on the current sequence, the researchers proposed BiLSTM. Compared with LSTM, BiLSTM adds a circular path for reading sequence information backward, using the data to obtain information about positive and negative directions and calculating the current output. It captures all the forward and backward information of the entire sequence at each step, thus training the network more efficiently. The structure of BiLSTM is illustrated in Fig. 3.

III. Experimental Settings

A. Selection of Evaluation Indicators and Data Preprocessing

The operating state of smart electricity meters mainly includes three aspects: quality, fault, and operating life cycle. The influencing factors are shown in Fig. 4. To a certain extent, each indicator can be indirectly or directly reflected by the measurement data of voltage, current, harmonic frequency, electric power, etc. [31].

In this article, the smart electricity meter relevant data are exported from the power consumption data acquisition system, marketing system (SG186), and measurement of integrated production dispatching system (MDS) of State Grid Corporation of China. The linking and repeated data of systems are reintegrated. The relevant data of successive 10 days for each smart electricity meter are collected. On each day, the sampling period is set at 1 h. Thus, 240 sets of data are obtained as one indicator data of the smart meter.

In addition, there exist multiple indicator data for each smart electricity meter, including current phase angle, voltage phase angle, frequency fluctuation, power factor, power direction, current, voltage, power, electrical energy, etc. However, current phase angle, voltage phase angle, frequency fluctuation, power direction, and electrical energy will not be selected as declared for the following reasons. Phase angle difference can be calculated indirectly from voltage, current, and power. The transmission power change can reflect frequency fluctuation to a certain extent in the power system, and the frequency fluctuation is not considered. Besides, there are random variables in the external environment, including temperature, humidity, and other measurable test states [14], [15]. Considering the limitations of the data, the random variables are not considered in the example analysis.

The current, voltage, and power are selected as indicators to evaluate the state of smart electricity meters, making the evaluation more accessible. Also, the overheating condition, insulation degree, and working strength of smart electricity meters can be characterized by these indicators to a certain extent.

Regarding smart electricity meter fault judgment, the main research method is to judge through the data collected by the online information acquisition system [3], [14], [15]. At present, the quantitative and qualitative analysis of the power change, current, voltage, and other electric data before and after the failure of the electric energy meter is used to determine whether the failure occurs [31]. The fault types include the expressed quantity of electric energy is not equal to the sum of various rates, the reverse active power indication being greater than zero, and the electric energy meter is creeping. This method can also judge the severity of various types of faults. The above studies have proved that the indicators selection in this article is feasible.

Due to unexpected reasons, some sampling data of smart electricity meters are missing, and there also exist abnormal values. The problematic data will not benefit the training model and may even adversely affect the model training. The data quality and characteristic analysis are used to preanalyze the initial data, including the missing value analysis, outlier analysis, and consistency analysis. After this procedure, data preprocessing is carried out, including data cleaning, integration, conversion, and specification. Therefore, the postprocessing sampling data form a dataset for the use of the deep learning model. The number of available meters is 3999.

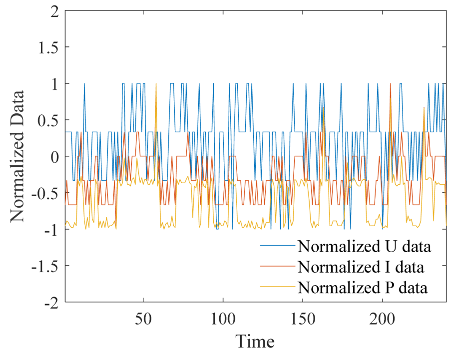

Fig. 5. Normalized indicator data of one smart electricity meter in a completely healthy state.

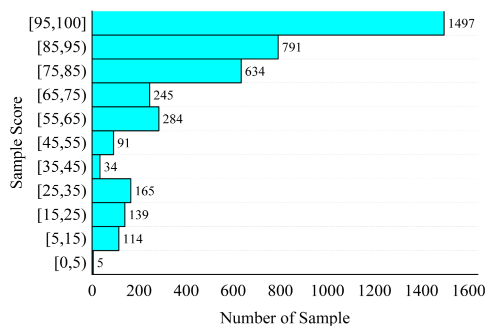

Fig. 6. Data distribution histogram of smart electricity meters based on different scores from expert system.

In addition, all smart electricity meters are automatically scored by the expert system, which has already been used in the State Grid Power Company. The expert system using the analytic hierarchy process considers the factors and indicators of smart meters as much as possible, referring to the manual analysis and judgment of the operating states of meters based on the actual and field data [33]. The scores of meters can serve as the labels for training the deep learning model in this experiment.

The evaluation model built in this experiment is a deep neural network, and the number of layers and neurons in the layer is not few. Besides, the amount of data is quite large. To make the network training converges rapidly, it is necessary to normalize and standardize the input data. The normalization formula is as follows:

$$x_t^* = 2 \times \left(\frac{x_t - \min\{x\}}{\max\{x\} - \min\{x\}} - 0.5 \right) \tag{7}$$

where x_t is data, $\min\{x\}$ is the minimum value in data, and $\max\{x\}$ is the maximum value in data, ${x^*}_t$ is the normalized data. The normalized data of one smart electricity meter in good condition is shown in Fig. 5.

The function of the model is pattern classification, and discrete scores must be converted into classification category labels. Therefore, the smart electricity meter's label has to be determined according to the score of the meter. To split a dataset by setting a fractional threshold, the data distribution should be first observed. Fig. 6 shows the data distribution histogram of smart electricity meters based on different scores from the expert system. In Fig. 6, most of the meters are in high scores (>75) because most of the meters are in good condition; after all, the minority belongs to unhealthy meters. The new labels are set as evenly distributed as possible to avoid the unbalanced distribution of meters' number on classifier's performance, as shown in Fig. 7. Fig. 7 classifies labels of 3999 m from high to low as "A," "B," "C," "D," "E," "F" with six degrees, where "A" indicates "completely healthy," "B" indicates "healthy," "C" indicates "good condition," "D" indicates the "general state," "E" indicates "state should be concerned," and "F" indicates "minor failure." Finally, the dataset contains 3999 groups of smart meters' sampling data. Among them, 3799 groups of sampling data are randomly selected to form a training set to train a deep learning model. The remaining 200 groups of sampling data form a test set. Simultaneously, the test set acts as a validation set for evaluating the learning performance of trained models.

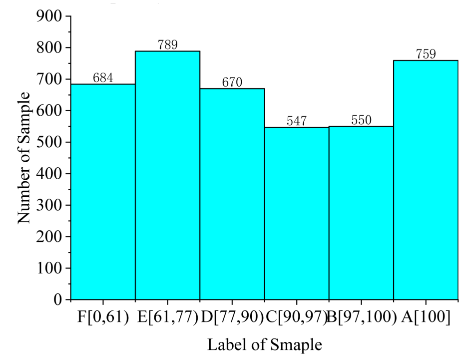

Fig. 7. Nearly evenly distributed labels for all smart electricity meter samples.

B. Construction of Deep Neutral Network

At present, neural networks play significant roles in many fields. In artificial intelligence applications, the data that need to be processed are often quite diverse. These complex and highly diverse data can be roughly divided into four categories: image, sequence, graph, and table. For each type of data, there is a corresponding neural network structure suitable for processing it.

In this article, the dataset used in the experiment includes long sequences, and there are time dimensions components and the static electrical element. If the convolution layer is used to build the network, multilayer 1-D convolution is needed to aggregate the sequences [34]. Nevertheless, using the sequence neural network can make the training network have some physical significance [35] because the sequence neural network can discover the relationship between the sequence changes with time and the operating state of the smart meters. Moreover, considering the model can be easily embedded into other large-scale systems in the future, it is necessary to select a simple and effective sequence neural network.

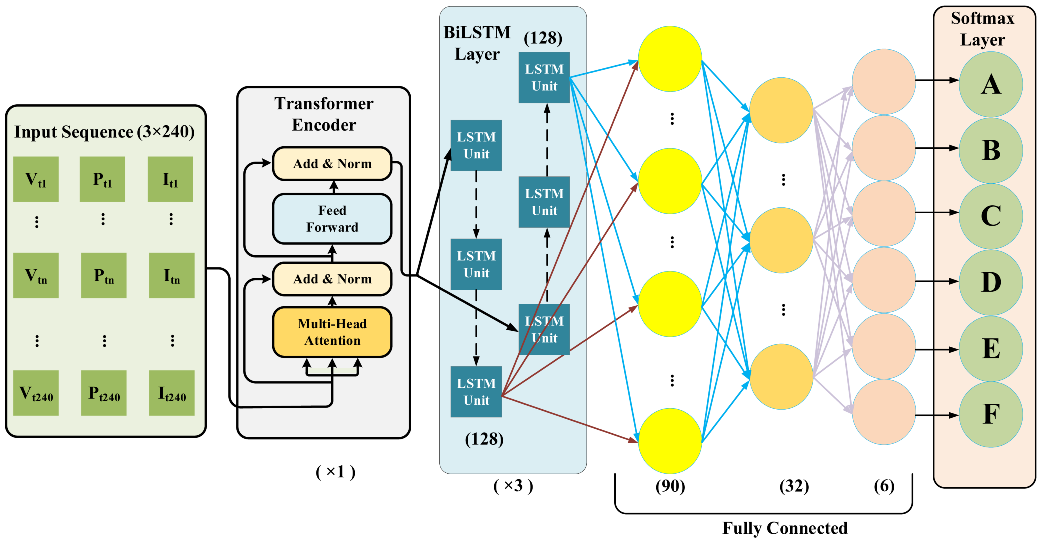

Fig. 8. Structure of the proposed transformer–encoder–BiLSTM network.

TABLE I
TRANSFORMER–ENCODER–BILSTM NETWORK STRUCTURE

	Type	Input size	Output size
1	Transformer-encoder layer	Batch size×240×3	Batch size×240×3
2	BiLSTM Layer 1	240×Batch size×3	240×Batch size×256
3	BiLSTM Layer 2	240×Batch size×256	240×Batch size×256
4	BiLSTM Layer 3	240×Batch size×256	Batch size×256
5	Fully connected layer 1	Batch size×256	Batch size×90
6	Fully connected layer 2	Batch size×90	Batch size×32
7	Fully connected layer 3	Batch size×32	Batch size×6
8	Softmax Layer	Batch size×6	Batch size×6

Based on the above analysis, this article uses transformer–encoder layer to extract the input sequence by attention mechanism. On the other hand, this layer has a certain denoising effect on the data. Then, BiLSTM layer replaces the traditional position encoding to extract the information of the time series dimension, and finally, fully connected layers are built for classification. The transformer–encoder–BiLSTM network, which is built in the experiment, consists of eight layers, and the network architecture is shown in Fig. 8. The input and output sizes are shown in Table I.

C. Training Parameter Settings for Deep Learning Network

Due to the high computer hardware requirement for deep learning model training, a server is used for calculation. The central processing unit (CPU) is used for preprocessing and loading data and the graphic processing unit (GPU) is used for processing large-scale matrices. The configuration of the server is shown in Table II.

The detailed architecture of the transformer–encoder–Bi LSTM and the training details are in the code given on GitHub.

TABLE II
SERVER HARDWARE AND SOFTWARE CONFIGURATION

Device	Model
CPU	Inter(R) Xeon(R) Gold 6268CL × 2
GPU	NVIDIA RTX A4000
RAM	128G
SOFTWARA	Python
PACKAGE	Numpy and Pytorch

TABLE III
STATISTICS RESULTS UNDER DIFFERENT ACTIVATION FUNCTIONS

Type	Average acc	Minimum epoch (acc≥90%)	Training times (acc>90)
selu	**99.5%**(98.5%–100%)	**794**	**20**
sigmoid	35% (23%–85%)	None	None
relu	53% (43%–96%)	2749	3
leaky relu	65% (44%–97%)	2505	7

IV. Result Analysis of Experiment

A. Transformer–Encoder–BiLSTM Performance Under Different Hyperparameters Settings

When training the transformer–encoder–BiLSTM network, this article found that the selection of activation function significantly impacts the speed and accuracy of training. Train the transformer—encoder–BiLSTM under *selu, sigmoid, relu, and leaky relu* in 20 repeated experiments, and the training results are given. The results are shown in Table III.

In Table III, the epoch of all experiments is 3000. Average acc represents the average accuracy of the test set under 20 repeated experiments, and the data in brackets represent the worst and best values in all experiments. The value of minimum epoch represents the value of epoch in which the accuracy of the test set reaches 90% at the earliest in 20 experiments. Training times

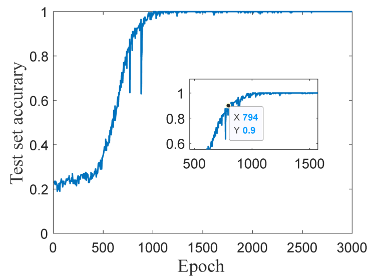

Fig. 9. Curve of test set accuracy of transformer–encoder–BiLSTM (best-trained one).

TABLE IV
STATISTICS RESULTS UNDER DIFFERENT RNN STRUCTURES

Type	Average acc
Transformer–encoder–BiLSTM	**99.5%** (98.5%–100%)
Transformer–encoder–LSTM	96% (93%–100%)
Transformer–encoder–BGRU	93% (87.5%–96.5%)
Transformer–encoder–GRU	91% (85%–93.5%)
Transformer–encoder–BRNN	80.5% (44%–93.5%)
Transformer–encoder–RNN	78.5% (44%–89.5%)

TABLE V
STATISTICAL RESULTS UNDER BiLSTM WITH DIFFERENT LAYERS

Type	Average acc	Minimum epoch (acc$\geqslant$90%)
One-layer BiLSTM	94.5%	501
Two-layer BiLSTM	97%	603
Three-layer BiLSTM	**99.5%**	**794**
Four-layer BiLSTM	99.5%	1021
Five-layer BiLSTM	99.5%	1161

represent the number of times the test set accuracy is more than 90% in the experiment. Table III shows that if the activation function uses *selu*, it can significantly improve the training efficiency. If the *sigmoid* function is used, the network will not achieve the desired effect in most cases. Fig. 9 is an accuracy curve of the test set belonging to the best-trained network in the experiments.

In the experiment, different RNN structures also affect the results. Table IV represents the test set accuracy under different structures in 20 repeated experiments.

It can be seen from Table IV that using a relatively complex BiLSTM layer can improve the network performance.

In the different experiments, the number of BiLSTM layers affects the performance. Table V shows the statistical results in 20 repeated experiments under BiLSTM with different layers. The experimental results show that the average accuracy of the test set decreases by 0.5%–3% for each less BiLSTM layer. Each additional BiLSTM layer will increase the minimum epoch by 100–400. According to the influence of training speed and network performance, the three-layer BiLSTM is the optimal choice of transformer–encoder–BiLSTM in Table V. There is no

TABLE VI
ROBUSTNESS ANALYSIS UNDER DIFFERENT TRAINED NETWORKS

Type	Average acc (test set without noise)	Average acc (test set with noise)
Transformer–encoder–BiLSTM	99.5% (98.5%–100%)	**89%** (84%–91%)
Transformer–encoder–LSTM	96% (93%–100%)	77% (73%–79%)
BiLSTM	73% (53%–98.5%)	43% (37%–55%)
LSTM	68% (44%–90.5%)	41.5% (33%–45%)

need for more layers of transformer–encoder from the current accuracy because adding more layers of transformer–encoder greatly increases the training time.

In addition, dropout must be used in the fully connected layer of transformer–encoder–BiLSTM. Otherwise, it will cause a 3%–5% performance loss.

B. Effect of Measurement Error on the Experimental Results of Various Methods

In practical application, it is considered that there will be measurement error ($<\pm 2\%$, maximum measurement error of smart electricity meter) when smart electricity meters record actual data. Therefore, this article adds 0.05 Gaussian noise to the normalized data of the test set to explore the robustness of the transformer—encoder–BiLSTM. Under 20 repeated experiments, the test set with noise is tested with the different trained networks (best-trained). The experimental results are shown in Table VI.

In Table VI, the average acc (test set without noise) represents the average accuracy of the test set under 20 repeated experiments. Average acc (test set with noise) represents the average accuracy of the network (best-trained one) in the test set with noise under 20 repeated experiments. It can be seen from Table VI that transformer–encoder–BiLSTM has stronger immunity to noise than other network structures. It demonstrates that transformer–encoder–BiLSTM has strong robustness. On the other hand, it also shows that this network has strong generalization performance for different data. The essence of anti-noise of the neural network is that each network structure is similar to a digital filter and has a certain denoising effect.

Moreover, additional experiments also found that increasing the layers of transformer–encoder and BiLSTM can improve the performance on the test set with noise by 0.5%–2%. However, it can greatly increase the training time requirements. In order to reach the minimum epoch, experiments show that 100–500 epochs are required for each additional layer of transformer–encoder and 100–400 epochs are required for each additional layer of BiLSTM. For instance, Fig. 10 is a test set accuracy curve of transformer–encoder–BiLSTM with five layers of BiLSTM, whose minimum epoch is 1161 and its average acc (test set with noise) is 90.5%. Therefore, the number of layers of transformer–encoder and BiLSTM can be selected according to the noise level of the actual data.

Fig. 11 is a loss function curve of network architecture in Table VI. Although in the experiment, the best trained BiLSTM

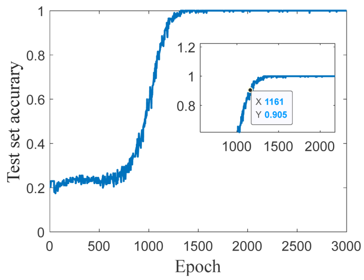

Fig. 10. Curve of the test set accuracy of transformer–encoder–BiLSTM with five layers of BiLSTM.

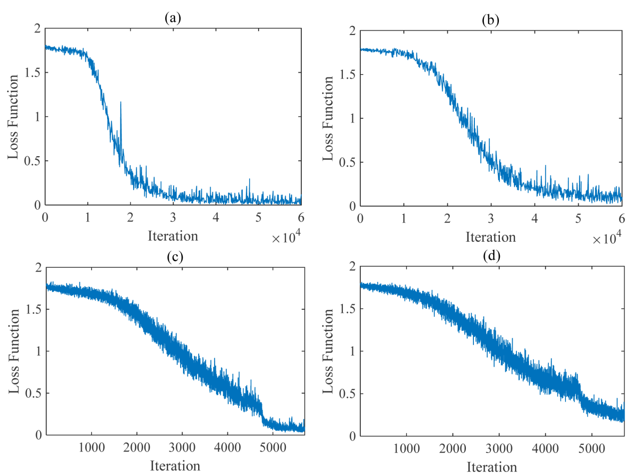

Fig. 11. Curve of loss function. (a) Transformer–encoder–BiLSTM (best-trained one). (b) Transformer–encoder–LSTM (best-trained one). (c) BiLSTM (best one). (d) LSTM (best-trained one).

(98.5%) and LSTM (90.5%) can achieve more than 90% accuracy on the test set, it requires researchers to make a lot of fine-tuning of each hyperparameter of the network, which is not conducive to other researchers to reproduce the results.

The measurement error of smart electricity meters is inevitable in the actual process. The accuracy of BiLSTM and LSTM in the test set with noise is much lower than that of transformer–encoder–BiLSTM and transformer–encoder–LSTM. If the BiLSTM and LSTM are used on the smart electricity meters with large measurement errors, it is unacceptable from the experimental results.

It can be seen from Fig. 11 that although BiLSTM and LSTM are not acceptable on the test set with noise, BiLSTM and LSTM have fewer iterations and can use less time to master some information of classification results before building a large network.

In the best-trained BiLSTM, only three samples' states are classified incorrectly using BiLSTM networks. The three misclassified samples are extracted for further observation to analyze the BiLSTM's classification performance in detail. Table VII presents prediction information of the BiLSTM regarding the misclassified samples.

In Table VII, the BiLSTM misclassifies the degree of sample 2 as degree D. However, this sample's score is 75.23, which is very close to the score threshold (77) between E and D. Thus, the misclassification of sample 2 is confusing. It is not typical for the model to misclassify this sample because it is between the degrees E and D. The model's decision-making habit tends to overestimate samples. Generally speaking, the predicted state degree of the model is higher than the actual situation of the sample. As can be seen from Table VII, the prediction labels of all three misclassified samples are better than their actual labels. This phenomenon can be explained as follows. Considering the dataset distribution, a large part of the dataset falls into the scope of high segments. Thus, it is not difficult to figure out that many samples dominate the weight adjustment of the neural network model with high degree labels, which makes the weight of the BiLSTM tend to be assimilated with these good label samples. The decision-making strategy of the network model tends to the good label.

TABLE VII
PREDICTION INFORMATION OF DEEP LEARNING NETWORK REGARDING THE MISCLASSIFIED SAMPLES

Wrong sample number	Sample score	Real label of the sample	Prediction label of the sample
1	76.11	E	A
2	75.23	E	D
3	21.92	F	D

TABLE VIII
STATISTICAL RESULTS OF MODEL WITH CNN ACCURACY

Type	Average acc	Average acc (test set with noise)
CNN (five layers)	83.5% (80%–85%)	58% (45%–65%)
CNN (five layers + residual connection)	87% (80%–93%)	68% (60.5%–74%)
CNN (two layers)	53% (45%–56%)	37% (22%–48%)
CNN–BiLSTM	90.5% (85%–96%)	74% (70%–79%)

It shows that when the samples are labeled, it should make the distribution of the test set uniform and consider making its score distribution uniform in some aspects. This phenomenon also reflects the lack of a large number of low score data in the dataset in this experiment. On the other hand, it also shows that transformer–encoder–BiLSTM has strong generalization ability and can overcome this problem.

C. Comparison With Convolutional Neural Network

In recent years, many researchers have used 1-D convolution to process sequence data [36], [37]. Therefore, four models based on a 1-D convolution module are constructed. They are CNN (five layers), CNN (five layers + residual connection), CNN (two layers), and CNN–BiLSTM. Under 20 repeated experiments, the average accuracy of the four models on the test set and test set with noise are shown in Table VIII. It can be seen from Table VIII that the accuracy of using a 1-D convolution neural network model is low, which is attributed to the weak expression ability of model, and the gradient dispersion and explosion caused by multilayer 1-D convolution network.

TABLE IX
PERFORMANCE COMPARISON WITH TRADITIONAL MACHINE LEARNING CLASSIFICATION MODEL

	Methods	Average acc (test set without noise)	Average computing time
Conventional machine learning (CPU)	Naive Bayes	22.5%	6.025 s
	Tree	25%	29.83 s
	Discriminant	30%	8.16 s
	KNN	31%	1 min 45 s
	Ensemble	35%	2 min 4 s
	SVM	37%	2 min 10 s
Conventional deep learning (GPU)	BP (10 layers)	64% (53%–77%)	7 min 52 s
	RNN	37.5% (25%–43.5%)	15 min 52 s
	BRNN	43.5% (34%–56%)	20 min 37 s
	GRU	53.5% (47%–57%)	19 min 37 s
	BGRU	67.5% (54.5%–83.5%)	25 min 14 s

TABLE X
PERFORMANCE COMPARISON WITH STATE-OF-THE-ART METHODS

Methods	Accuracy	Intelligence	Robust to measurement error	Rapidity	Applicability in each meter
Proposed method	>95%	√	√	√	√
[17]	Not mentioned	×	Not mentioned	×	√
[18]	Not mentioned	×	Not mentioned	×	√
[24]	>95%	√	×	√	Standard meter
[25]	Not mentioned	×	×	×	√
[26]	>95%	×	×	×	√
[27]	Not mentioned	√	×	×	√

D. Comparison With Other Traditional Machine Learning

In addition to the deep learning model mentioned above, the classification model based on transformer–encoder–BiLSTM is also compared with traditional machine learning (computing using CPU) and deep learning methods (computing using GPU). The performance comparison result of proposed methods with other models is shown in Table IX. It can be seen that the accuracy index and computing time index are quite different, and the accuracy seems to be proportional to the computing time. Generally, the stability of the machine learning model is high, while the randomness of deep learning is high, so 20 repeated experiments are done for each deep learning model.

From Table IX, it can be seen that the following conditions hold:

1) The machine learning model does not have a more complex nonlinear relationship, so the calculation speed is faster than the deep learning model. The accuracy of the machine learning model in this experiment is relatively low. The accuracy is closely related to the structure of data that the machine learning model is suitable for processing. Machine learning is not ideal for processing large sample sizes or high-dimensional data. In this experiment, the number of smart electricity meters is 3999. Thus, the accuracy of machine learning is low.
2) In the deep learning model, dropout and regularization make some neurons disappear, and the weights are decreased to achieve automatic feature extraction. Deep learning has a more complex nonlinear relationship than manual feature extraction in machine learning. Thus, the relationship between indicators and labels can be better mined.
3) Deep learning models have better performance than traditional machine learning models, revealing the robustness of deep learning networks in feature extraction of large-scale datasets of smart electricity meters. Deep neural network needs large-scale data to mine the relationship between inputs and outputs. It also requires much computational power to train the network to get the appropriate functional relationship [34]–[38]. Traditional machine learning methods are suitable for small-scale datasets [39], [40]. Machine learning is not easy to construct complex functional relationships between inputs and outputs. Besides, its performance largely depends on the number and rationality of selected indicators [38].

In conclusion, the new architecture of transformer–encoder–BiLSTM proposed in this article is far better than any of the above models with stronger generalization ability and greater robustness. Accordingly, much computing power and time are spent on training this model.

E. Comparison With Traditional Evaluation Methods

Because the characteristic of input indicators for other state evaluation methods are different from that of the proposed method and the limitations of data, this article only investigates some relevant methods according to [17], [18], [24]–[27] in the introduction, and the performance comparison with state-of-the-art methods are shown in Table X. It can be seen from Table X that compared with other methods, the proposed method has certain advantages in accuracy, resistance to measurement error, rapidity, intelligence, and applicability, which is related to the advantages of using the time-series data and deep learning structure.

V. SUGGESTION AND FUTURE ANALYSIS

A host of methods are currently proposing to evaluate the operating state through the physical condition of the smart electricity meter [19], [20]. This article finds a suitable deep learning method for state evaluation and lays a solid foundation for integrating the proposed model into the online system. The proposed state evaluation method is accurate, and it can provide an effective method for power companies. However, the following suggestions are given:

1) Among the collected data of smart electricity meters, several electric quantity data are selected as indicators in this article. With the update of the smart electricity meter's data, more indicators can be used in the future.
2) The data of each smart electricity meter can be sampled and collected in every time step for the future online system. The operating state evaluation for smart electricity meters can be trigged and realized every other month.
3) Due to the acquisition accuracies of variable smart meters being similar, smart meters' data of different regions have a similar distribution. Thus, transfer learning is suitable to generalize the model, for instance, from A Power Grid to

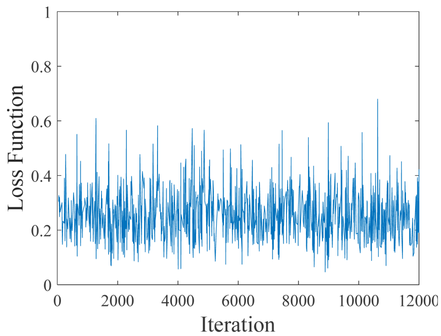

Fig. 12. Loss function curve using fine-tune.

B Power Grid. Before applying the trained model, it is best to collect a certain amount of data from corresponding regions to fine-tune the model's parameters. Fine-tune can make the model more consistent with the smart electricity meters in the new region and improve the robustness of the model. On the other hand, using the previously trained model can speed up the training process. For example, 300 training data are randomly selected, and then 0.2 Gaussian noise is added to the extracted normalized data to simulate the sequence information of smart electricity meters in other regions. Then, we fine-tune the model's parameters. Fig. 12 is the curve of the loss function. Compared with Fig. 11(a), it can be seen that this strategy greatly increases the training speed of the network.

4) The trained model can be updated. According to different customer consumption profiles, fine-tune and other techniques (for instance, lifelong learning) can train the previously trained model. Besides, the elastic weight consolidation [41], incremental classifier and representation learning [42] techniques can be used to add new operating states or new tasks to the previously trained model.

In the future, this proposed method will have a considerable development prospect of the online application when it is combined with the Internet of things, cloud computing, and other new technologies. It can provide an automatic evaluation process for stakeholders by replacing the manual sampling and scoring method. Besides, with the continuous development of algorithms, computing power, and big data, which are three important factors, affecting deep learning performance [43], the proposed method with more strong generalization ability can better mine the relationship between the operating state of smart electricity meters and electrical quantities. It can be easily integrated into the MDS system of state grid to help the operation personnel manage meters.

VI. Conclusion

This article establishes and facilitates an evaluation method of the operating state for smart electricity meters based on smart electricity meter monitoring data. First, this article analyzes and selects the evaluation indicators. Then, a scheme of mining the state information of smart electricity meters using transformer–encoder–BiLSTM is constructed. Third, the experimental verification is performed based on groups of smart electricity meters' data; the effects of different hyperparameters and measurement error on transformer–encoder–BiLSTM performance are discussed; the article compares the performance of the transformer–encoder–BiLSTM-based classifier with other models. Finally, the suggestions and future analyses of using the proposed method are provided. The conclusions are as follows:

1) Compared with the traditional expert system, the proposed model is more intelligent. The evaluation process does not require much workforce, human intervention, and material resources. The proposed method does not need to extract the features artificially, and it can realize automatic extraction. Besides, although the proposed method uses the scores from the expert system for training the deep learning model, there will be no need for expert scoring results in the future after the model is well trained.
2) The experiment shows the classification average accuracy of the transformer–encoder–BiLSTM-based model can be as high as 99.5%. The proposed model can make an accurate evaluation of the operating states of smart electricity meters. The evaluation result can provide a reference for the intellectual maintenance and replacement of smart electricity meters.
3) The comparison study shows that the new architecture of transformer–encoder–BiLSTM has stronger generalization ability and greater robustness to noise. Compared with the traditional machine learning method, although the proposed method increases the training time of the network, the accuracy has been significantly improved, and has specified reliability.

References

[1] T. Sirojan, S. Lu, B. T. Phung, and E. Ambikairajah, "Embedded edge computing for real-time smart meter data analytics," in *Proc. Int. Conf. Smart Energy Syst. Technol.*, 2019, pp. 1–5.
[2] Q. Malik, A. Zia, R. Ahmad, M. A. Butt, and Z. A. Javed, "Design and operation of smart energy meter for effective energy utilization in smart cities," in *Proc. IEEE Conf. Sustain. Utilization Develop. Eng. Technol.*, 2019, pp. 219–223.
[3] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, "Smart meter data analytics for distribution network connectivity verification," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1964–1971, Jul. 2015.
[4] A. Pasdar and S. Mirzakuchaki, "A solution to remote detecting of illegal electricity usage based on smart metering," in *Proc. 2nd Int. Workshop Soft Comput. Appl.*, 2007, pp. 163–167.
[5] D. George and G. L. Swan, "A method for distinguishing appliance, lighting and plug load profiles from electricity 'smart meter' datasets," *Energy Buildings*, vol. 134, pp. 212–222, 2017.
[6] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3125–3148, May 2019.
[7] P. Carroll, T. Murphy, and M. Hanley, "Household classification using smart meter data," *J. Offical Statist.*, vol. 34, no. 1, pp. 1–25, 2018.
[8] A. N. Funde, M. M. Dhabu, and A. Paramasivam, "Motif-based association rule mining and clustering technique for determining energy usage patterns for smart meter data," *Sustain. Cities Soc.*, vol. 46, 2019, Art. no. 101415.
[9] E. Ebeid, R. Heick, and H. R. Jacobsen, "Deducing energy consumer behavior from smart meter data," *Future Internet*, vol. 9, 2017, Art. no. 293.
[10] A. Tureczek, S. P. Nielsen, and H. Madsen, "Electricity consumption clustering using smart meter data," *Energies*, vol. 11, 2018, Art. no. 8594.
[11] A. M. Alonso, F. J. Nogales, and C. Ruiz, "Hierarchical clustering for smart meter electricity loads based on quantile autocovariances," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4522–4530, Sep. 2020.

[12] L. Yang, X. Chen, J. Zhang, and H. V. Poor, "Cost-effective and privacy-preserving energy management for smart meters," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 486–495, Jan. 2015.

[13] K. Zheng, Q. Chen, Y. Wang, C. Kang, and Q. Xia, "A novel combined data-driven approach for electricity theft detection," *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1809–1819, Mar. 2019.

[14] F. Liu, C. Liang, Q. He, L. Wang, C. Huang, and S. Hu, "An approach for online smart meter error estimation," in *Proc. Conf. Precis. Electromagn. Meas.*, 2020, pp. 1–2.

[15] F. Liu, C. Liang, and Q. He, "Remote malfunctional smart meter detection in edge computing environment," *IEEE Access*, vol. 8, pp. 67436–67443, 2020.

[16] Y. Gao, B. Foggo, and N. Yu, "A physically inspired data-driven model for electricity theft detection with smart meter data," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 5076–5088, Sep. 2019.

[17] X. Cheng, M. Yu, M. Liu, R. Huang, L. Xie, and H. Tan, "Research on comprehensive performance evaluation method of smart energy meter," in *Proc. 3rd Int. Conf. Mech., Control Comput. Eng.*, 2018, pp. 450–455.

[18] C. Y. Ying, D. Jie, Z. Feng, X. Ji, and Y. H. Xiao, "Application of variable weight fuzzy analytic hierarchy process in evaluation of electric energy meter," in *Proc. IEEE 2nd Adv. Inf. Technol., Electron. Automat. Control Conf.*, 2017, pp. 1985–1989.

[19] J. Q. Zhang, S. Guo, G. Li, and Y. Chen, "Reliability evaluation model of intelligent watt hour meter based on multiple big data fusion," *Elect. Meas. Instrum.*, vol. 1, no. 7, pp. 1–7, 2022.

[20] D. Xu, Q. Wei, E. A. Elsayed, Y. Chen, and R. Kang, "Multivariate degradation modeling of smart electricity meter with multiple performance characteristics via vine copulas," *Qual. Rel. Eng.*, vol. 33, no. 4, pp. 803–821, 2017.

[21] W. Luan, J. Peng, M. Maras, J. Lo, and B. Harapnuk, "Smart meter data analytics for distribution network connectivity verification," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1964–1971, Jul. 2015.

[22] C. Wei, Z. Bing, F. Zhancheng, and F. Yilun, "The research of smart electricity meter whole performance automatic detection technology," in *Proc. IEEE Int. Conf. Comput. Sci. Automat. Eng.*, 2012, pp. 431–434.

[23] Z. Feng *et al.*, "Technology and application of multidimensional remote monitoring system for electric energy meter based on decision tree group," in *Proc. IEEE 3rd Adv. Inf. Manage., Commun., Electron. Automat. Control Conf.*, 2019, pp. 1141–1145.

[24] Y. Jiao, H. Li, C. Hu, Z. Zhang, and C. Zhang, "Data-driven evaluation for error states of standard electricity meters on automatic verification assembly line," *IEEE Trans. Ind. Informat.*, vol. 15, no. 9, pp. 4999–5010, Sep. 2019.

[25] S. Yip, K. Wong, and W. Hew, "Detection of energy theft and defective smart meters in smart grids using linear regression," *Int. J. Elect. Power Energy Syst.*, vol. 91, pp. 230–240, Oct. 2017.

[26] L. Helong, Y. Haibo, and Y. Jinshuai, "Intelligent energy meter fault prediction based on machine learning," in *Proc. 15th Int. Conf. Comput. Intell. Secur.*, 2019, pp. 296–300.

[27] Z. Feng *et al.*, "Construction of multidimensional electric energy meter abnormal diagnosis model based on decision tree group," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf.*, 2019, pp. 1687–1691.

[28] A. Vaswani *et al.*, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.

[29] C. Subakan, M. Ravanelli, S. Cornell, M. Bronzi, and J. Zhong, "Attention is all you need in speech separation," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2021, pp. 21–25.

[30] J. Liu, G. Wang, L. Duan, K. Abdiyeva, and A. C. Kot, "Skeleton-based human action recognition with global context-aware attention LSTM networks," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1586–1599, Apr. 2018.

[31] W. Xi, Y. Ji, and J. Zhang, "Review on the evaluation technology of smart watt hour meter operation state," *Elect. Meas. Instrum.*, vol. 57, no. 3, pp. 134–141, 2020..

[32] J. Fan, X. Chen, and Y. Zhou, "An intelligent analytical method of ab-normal metering device based on power consumption information collection system," *Elect. Meas. Instrum.*, vol. 50, no. 11, pp. 4–9, 2013.

[33] B. Wang, J. Yin, S. S. Hu, Y. Luo, Y. Pan, and X. Xiao, "Research on reliability distribution technology of intelligent watt hour meter based on analytic hierarchy process and group decision," *Elect. Meas. Instrum.*, vol. 58, no. 12, pp. 169–174, 2021.

[34] H. Kameoka, K. Tanaka, D. Kwaśny, T. Kaneko, and N. Hojo, "ConvS2S-VC: Fully convolutional sequence-to-sequence voice conversion," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 28, pp. 1849–1863, 2020.

[35] Q. Yang, Z. He, F. Ge, and Y. Zhang, "Sequence-to-sequence prediction of personal computer software by recurrent neural network," in *Proc. Int. Joint Conf. Neural Netw.*, 2017, pp. 934–940.

[36] L. Jiao, H. Wu, H. Wang, and R. Bie, "Text recovery via deep CNN-BiLSTM recognition and Bayesian inference," *IEEE Access*, vol. 6, pp. 76416–76428, 2018.

[37] J.-H. Jeong, K.-H. Shim, D.-J. Kim, and S.-W. Lee, "Brain-controlled robotic arm system based on multi-directional CNN-BiLSTM network using EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1226–1238, May 2020.

[38] M. Kim, S. Lee, and J. Kim, "A wide & deep learning sharing input data for regression analysis," in *Proc. IEEE Int. Conf. Big Data Smart Comput.*, 2020, pp. 8–12.

[39] A. Glowacz, "Ventilation diagnosis of angle grinder using thermal imaging," *Sensors*, vol. 21, 2021, Art. no. 2853.

[40] B. Seetharamulu, B. N. K. Reddy, and K. B. Naidu, "Deep learning for sentiment analysis based on customer reviews," in *Proc. 11th Int. Conf. Comput., Commun. Netw. Technol.*, 2020, pp. 1–5.

[41] B. Maschler, H. Vietz, N. Jazdi, and M. Weyrich, "Continual learning of fault prediction for turbofan engines using deep learning with elastic weight consolidation," in *Proc. 25th IEEE Int. Conf. Emerg. Technol. Factory Automat.*, 2020, pp. 959–966.

[42] S. Rebuffi, A. Kolesnikov, G. Sperl, and C. H. Lampert, "iCaRL: Incremental classifier and representation learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5533–5542.

[43] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," *CSEE J. Power Energy Syst.*, vol. 7, no. 2, pp. 209–220, Mar. 2021.

Zhongyong Zhao (Member, IEEE) was born in Guangyuan, China. He received the B.S. and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 2011 and 2017, respectively.

He is currently an Associate Professor with the College of Engineering and Technology, Southwest University, Chongqing. His research interests include condition monitoring and fault diagnosing for HV apparatus, and artificial intelligence.

Dr. Zhao was the recipient of the scholarship from the China Scholarship Council to enable him to attend a joint-training Ph.D. program in Curtin University, Perth, WA, Australia, in 2015–2016.

Yu Chen was born in Wenzhou, China, in 2000. He is currently working toward the bachelor's degree in electrical engineering with the College of Engineering and Technology, Southwest University, Chongqing, China.

His research interests include condition monitoring and fault diagnosing for power devices and application of artificial intelligence.

Jiangnan Liu was born in Ganzhou, China, in 1999. He received the B.Sc. degree in electrical engineering from the College of Engineering and Technology, Southwest University, Chongqing, China, in 2020. He is currently working toward the master's degree in electrical engineering with the Chongqing University, Chongqing.

His research interests include condition monitoring and fault diagnosing for HV apparatus and pulsed power technology.

Yingying Cheng was born in Nanchong, China, in 1976. She received the B.S. and M.S. degrees in electrical engineering from Chongqing University, Chongqing, China, in 1997 and 2000, respectively.

She is currently a Professor of Engineering. Her research interests include electric energy measurement, test detection, and intelligent measurement.

Chenguo Yao (Member, IEEE) was born in Nanchong, China. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from Chongqing University, Chongqing, China, in 1997, 2000, and 2003, respectively.

He became a Professor with the School of Electrical Engineering, Chongqing University, in 2007. He was a Visiting Scholar with Dominion University, Norfolk, VA, USA, from 2017 to 2018.

Chao Tang (Member, IEEE) was born in Sichuan, China, in 1981. He received the M.S. and Ph.D. degrees in electrical engineering from the Chongqing University, Chongqing, China, in 2007 and 2010, respectively.

As a Ph.D. student (2008–2009) and as a Visiting Scholar (2013–2013 and 2015–2016), he studied with the Tony Davies High Voltage Laboratory, University of Southampton, Southampton, U.K., doing some researches on the dielectric response characteristics and space charge behaviors of oil-paper insulation. He is currently a Professor with the College of Engineering Technology, Southwest University, Chongqing.