B20041231 章春阳
A Novel Deep Neural Network Framework for State Evaluation and Fault Diagnosis in Distribution Station

*Note: Sub-titles are not captured in Xplore and should not be used

1st B20041231 章春阳
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

摘要—With the growing demand for reliable power supply from smart grid infrastructures, the assessment of operational status and fault diagnosis in distribution substations is becoming increasingly important. Traditional methods often fall short in providing real-time and accurate analysis, necessitating advancements in intelligent monitoring systems. To address this gap, a deep neural network framework, called SFCED, is proposed specifically for real-time state evaluation and fault diagnosis in distribution substations. This framework uses signature matrices to capture and represent correlations within multivariate time-series data, enabling a comprehensive understanding of distribution substation dynamics. SFCED integrates a fully convolutional encoder-decoder architecture, enabling the extraction of deep features from data and accurate reconstruction of system states. The empirical results, obtained from extensive comparisons with conventional CNN and LSTM Encoder-Decoder models, confirm the effectiveness of SFCED, particularly in achieving a high balance between precision and recall. The research demonstrates the practical applicability of SFCED in real-world scenarios, offering significant improvements in maintenance and operation efficiency for intelligent distribution substations.

Index Terms—**State Evaluation, Distribution Station, Deep Learning, Signature Matrices, Fault Diagnosis**

Identify applicable funding agency here. If none, delete this.

I. Introduction

With the transformation of the global energy mix and the modernization of power systems, the development of smart grid technologies has become key to promoting sustainable energy management and efficient energy distribution [1]. Smart grids not only optimize energy consumption, but also enhance grid reliability and resilience [2]. Distribution substation is the core end node of the distribution grid, which is an important link to ensure the reliability and quality of power supply for low-voltage users [3]. Through the transformation of the intelligent auxiliary and visualization system of the distribution station house, data such as environmental and electrical quantities are uploaded using intelligent devices such as intelligent gateways, sensors, and visualization devices, and the data integration is initially completed to realize the functions of environmental monitoring, auxiliary equipment control, security monitoring, and so on [4]. However, it is still challenging to effectively process and deeply analyze the large amount of multi-source heterogeneous data collected from distribution substation in order to realize the condition assessment and fault diagnosis of distribution substation.

Currently, significant progress has been made in the research field of power system state evaluation and fault diagnosis, and mainly focuses on the single devices in distribution networks [5]. Research has covered a variety of methods, including fuzzy comprehensive evaluation [6], cloud model [7], Dempster-Shafer evidence theory [8] and deep learning [9]. Deep learning, in particular, has been recognized as having great potential for power system state evaluation and fault diagnosis due to its remarkable ability to process complex, high-dimensional data and extract potential features. Guo et al. [10] propose a comprehensive assessment method for the health index (HI) model of transformers. For the first time, the method considered the aging process, load and operating environment of the transformer insulation and combined with field test data in order to form a theoretical health index HI1 and a test health index HI2. Finally, the transformer's health is characterized by combining HI1 and HI2. Zhou et al. [11] propose a comprehensive health assessment method based on Dempster-Shafer (D-S) multi-evidence fusion for medium voltage switchgear. The method first establishes an independent mapping from the subsystem diagnostic parameters to the health index (HI) of the respective subsystems using a weighted summation method. Subsequently, the HI of the subsystems are integrated into the same sample space within the framework of the D-S methodology, and then combined into a composite HI for a comprehensive assessment of the switchgear health. Xing et al. [12] describe a multimodal cooperative neural network for health assessment of power transformers. This method uses a one-dimensional convolutional neural network to extract features from dissolved gas data and a deep residual squeeze-excitation network to extract features from infrared images. Experimental results show that this method achieves high classification accuracy in the health assessment of power transformers. Liu et al. [13] propose a new transformer fault diagnosis framework that utilizes kernel principal component analysis to extract features from dissolved gas analysis data and balances data from different fault types using the SMOTE+ENN method. This framework excels in transformer fault classification problems. Xu et al. [14] propose a perception method for monitoring the operational status of secondary equipment in distribution substation. This method is based on wavelet transform and ResNet. Firstly, wavelet transform is used to process noise in environmental and operational data. Then, data features are extracted through a time-frequency analysis method. Finally, the ResNet model is used to classify the time-frequency graphs and determine the current operating status of the distribution substaion.

However, although the above methods have made significant progress in state evaluation and fault diagnosis of specific equipment, they usually fail to comprehensively assess the operational status of distribution substations. With the increase in the number of distribution substation and equipment, the traditional state evaluation and fault diagnosis methods have gradually become burdensome and inefficient. In order to reduce the inspection workload of operation and maintenance personnel, detect and deal with the defects in distribution substations and their equipments in a timely manner, and carry out targeted maintenance, there is an urgent need to process and deeply analyze a large amount of data collected in distribution substations, such as environmental and electrical quantities, in order to realize real-time condition assessment and fault diagnosis of intelligent distribution substations.

Based on the above discussion, a novel deep neural network framework is proposed in this study, which is specifically used for the state assessment and fault diagnosis of distribution substation. The framework fully considers the multi-source data characteristics in the distribution substation and introduces signature matrices to effectively represent various state information. We adopt a fully convolutional encoder-decoder structure, which is capable of fully learning the deep features in the data, so as to accurately assess the operational state of the distribution substation and detect faults in a timely manner.

In this study, our main contributions include the following:

1) A novel deep neural network framework is specifically designed for state assessment and fault diagnosis of smart distribution substation.
2) Our approach takes into account the multi-source data characteristics within the distribution substation and introduces signature matrices, which effectively integrates and represents various state information and improves the efficiency and accuracy of data processing.
3) The framework in this study has been tested on real distribution station data, and the results demonstrate its effectiveness and practicality in condition assessment and fault detection.

II. Proposed Method

As shown in Fig. 1, the framework proposed in this study consists of two key components: the signature matrices and the fully convolutional encoder-decoder. Firstly, the signature matrices are responsible for extracting and characterizing state information at different scales from multi-source data of the distribution substation, effectively integrating the data from different sensors, including data such as environmental and electrical quantities. Secondly, fully convolutional encoder-decoder is used to learn these features in depth. Fully convolutional encoder is responsible for the learning of deep features, while fully convolutional decoder is used to reconstruct the signature matrices. In most cases, the residual signature matrices are utilized for state evaluation and fault diagnosis. Furthermore, by employing the squared loss function, our model is able to perform end-to-end learning and optimize the performance of the entire network.

A. Signature Matrices

In order to deeply analyze the intrinsic correlation between multivariate time-series data in the intelligent monitoring system of distribution substation, we introduce signature matrice. Its purpose is to characterize the interrelationships among various types of sensor data (e.g., temperature, humidity, current, voltage, etc.) in different time periods. For a given time period $t-w$ to t, we construct an $n \times n$ signature matrix M^t, where n represents the number of sensors in the distribution substation. This matrix is formed by calculating the pairwise inner product of different pairs of time series in that time period. The computational formula is as follows:

$$\mathbf{x}_i^w = \left(x_i^{t-w}, x_i^{t-(w-1)}, \cdots, x_i^t\right) \tag{1}$$

$$\mathbf{x}_j^w = \left(x_j^{t-w}, x_j^{t-(w-1)}, \cdots, x_j^t\right) \tag{2}$$

$$m_{ij}^t = \frac{\sum_{\delta=0}^{w} x_i^{t-\delta} x_j^{t-\delta}}{\kappa} \tag{3}$$

where κ is a rescale factor ($\kappa = w$). x_i^w and x_j^w represent two different time series in the multivariate time series

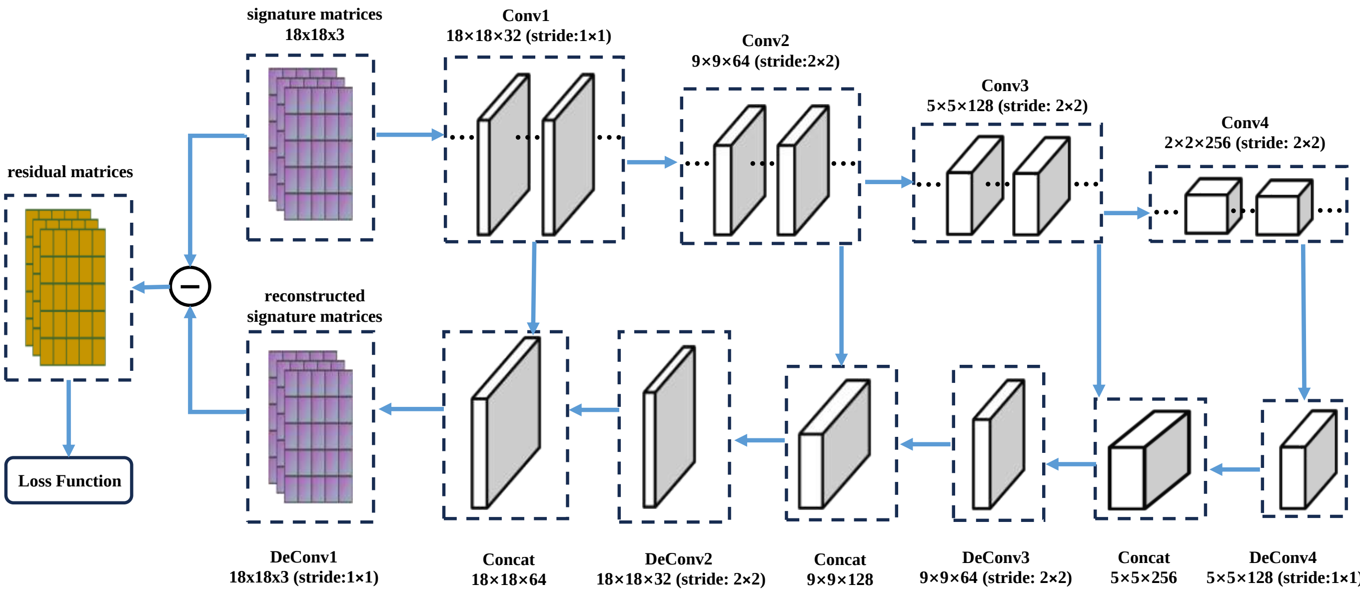


图 1: The framework of SFCED.

segment, and m_{ij}^t is the correlation between x_i^w and x_j^w. The interval between the two segments is set to 10. In addition, this study sets three signature matrices with different window lengths ($w = 5, 10, 15$) at each time step to take into account the dynamic characteristics of the system over multiple time scales.

B. Fully Convolutional Encoder

We use a fully convolutional encoder [15] to process and learn the spatial relationships present in the system's signature matrices. These matrices contain the system state across multiple dimensions and are combined into a composite tensor, symbolically represented as X_t^0 in the domain of $\mathbb{R}^{n\times n\times s}$, where s denotes the scale of the data under consideration. This tensor is then passed sequentially through several convolutional layers for feature extraction.

Let X_t^{l-1} denotes the extracted features from the previous layer, $l-1$, in the architecture. The convolutional operation in the l-th layer can be expressed mathematically as:

$$x_t^l = f\left(W^l * x_t^{l-1} + b^l\right), \tag{4}$$

where $*$ represents the convolutional operation and $f(\cdot)$ represents the activation function. The terms W^l and b^l correspond to the convolutional kernels and bias elements for each respective layer l. For the coding layers, we have adopted the Scaled Exponential Linear Unit (SELU) for activation due to its self-normalising properties. The encoder consists of four convolutional layers, designated Conv1 to Conv4, with 32 kernels of size $3\times3\times3$, 64 kernels of size $3\times3\times32$, 128 kernels of size $2\times2\times64$, and 256 kernels of size $2\times2\times128$, respectively.

C. Fully Convolutional Decoder

To decode the feature maps obtained in the previous step and obtain the reconstructed signature matrices, a convolutional decoder is designed as follows:

$$\hat{x}_t^{l-1} = \begin{cases} f\left(\hat{W}^l \oslash x_t^l + \hat{b}^l\right), & \text{for } l = 4 \\ f\left(\hat{W}^l \oslash \left(\hat{x}_t^l \oplus x_t^l\right) + \hat{b}^l\right), & \text{for } l = 3,2,1 \end{cases} \tag{5}$$

$\oslash$ represents the deconvolution operation, which enlarges the feature maps to their original dimensions. $\oplus$ is the concatenation operation of the feature maps from the equivalent encoder layers, which adds detailed spatial information back into the data. $f(\cdot)$ is the activation function, identical to that used in the encoder, which ensures coherence throughout the network. $\hat{W}^l$ and $\hat{b}^l$ are the learned weights and biases for the deconvolution layers at level l, which restore the feature map x_t^l to its previously encoded state x_t^{l+1}. The decoder consists of four deconvolutional layers, designated DeConv1 to DeConv4, with 3 kernels of size $3 \times 3 \times 64$, 32 kernels of size $2 \times 2 \times 128$, 64 kernels of size $2 \times 2 \times 256$, and 128 kernels of size $4 \times 4 \times 256$, respectively.

D. Distribution substation state evaluation and fault diagnosis process

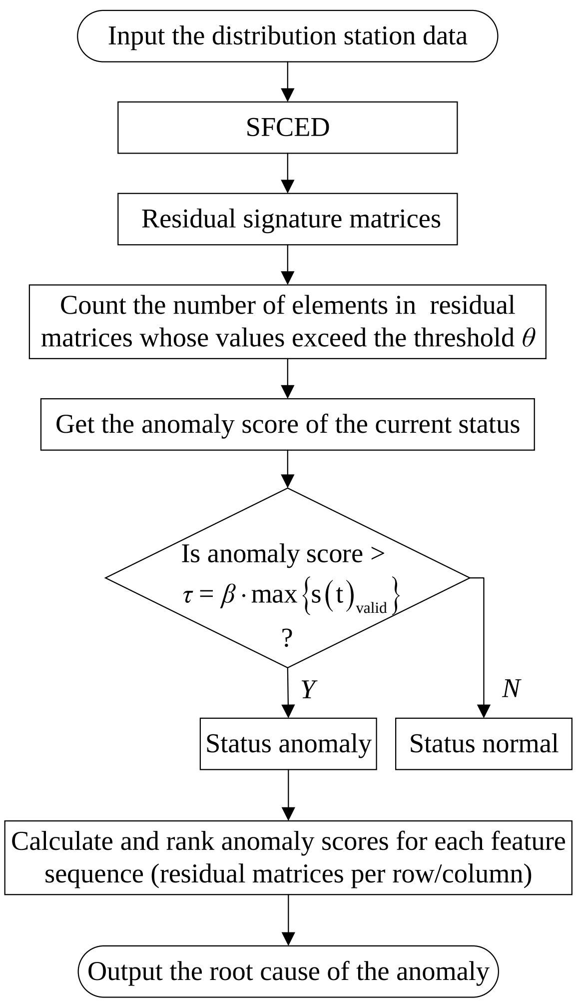


图 2: Flow chart of State Evaluation and Fault Diagnosis in Distribution Station.

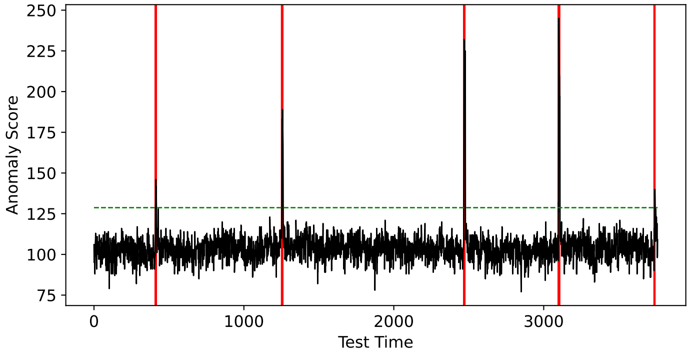
As shown in Fig. 2 The process begins with the input of data from the distribution station into the SFCED model. Once the data is received, the SFCED model processes it to generate residual signature matrices. Next, a crucial step involves counting the number of elements in the residual matrices that exceed a predefined threshold, denoted as θ. This threshold is empirically set based on historical data and expert domain knowledge. The count of these elements is then used to confirm the anomaly score for the current status of the distribution station. This anomaly score is then assessed against a dynamic threshold τ, which is determined by multiplying β (a scaling factor) with the maximum anomaly score observed during the validation period, $s(t)_{\text{valid}}$. If the anomaly score exceeds the dynamic threshold τ, the system identifies the current state as an anomaly; otherwise, it is considered normal. For anomalies detected, the next step is to diagnose the potential causes. This is done by calculating and ranking the anomaly scores for each feature sequence derived from individual rows/columns of the residual matrices. By ranking these scores, it is possible to ascertain the most likely sources of anomalies within the distribution station's operation.

III. Case Study

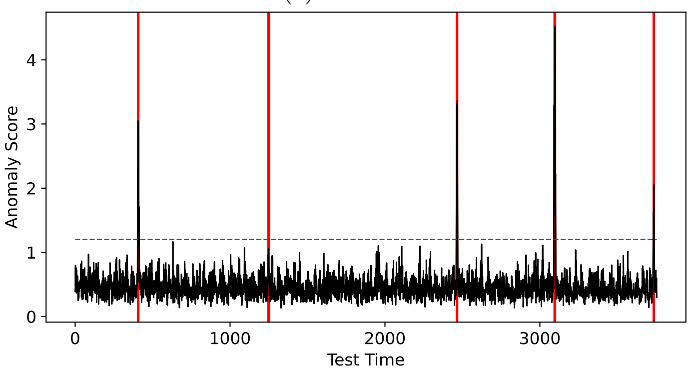
The dataset used in this study comes from the data collected by the Intelligent Auxiliary Monitoring System for a distribution substation in Nanjing, China, spanning the period from September 10, 2022 to September 10, 2023, a period of one year. The dataset covers more than 190 variables and is collected at an hourly frequency. To effectively assess the status of the station house and perform fault diagnosis, we selected 18 major variables from it for analysis. The dataset contains $365 * 24 = 8760$ data points and has 5 anomalies.

A. Evaluation Metrics

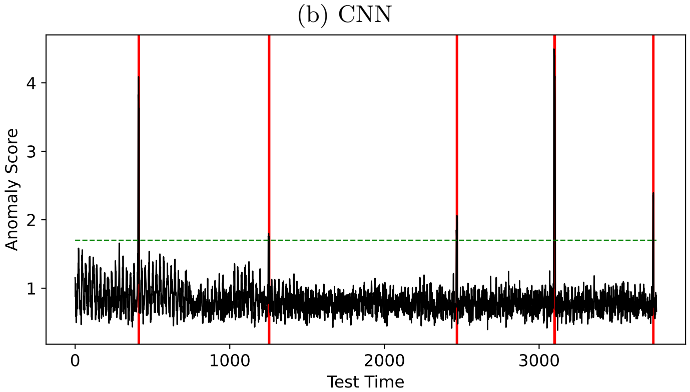
We use three metrics: Precision, Recall, and the F1 Score, to appraise the efficacy of proposed method. The determination of an optimal threshold for anomaly detection follows the expert's recommendation, and is computed as $\beta \cdot \max(s(t)_{valid})$, where $s(t)_{valid}$ denotes the validation period's anomaly scores, and β is tuned within the range [1, 2] to optimize the F1 Score [16].



(a) SFCED



(b) CNN



(c) LSTM Encoder-Decoder

图 3: Comparison of State Evaluation Results

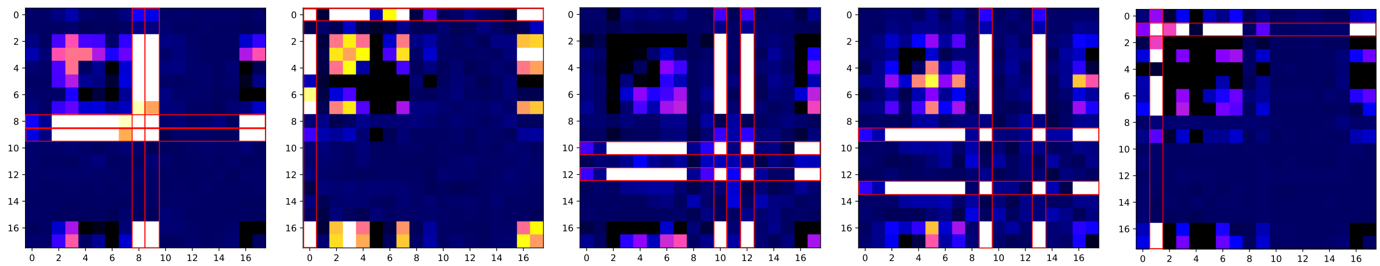
B. *State Evaluation and Fault Diagnosis in Distribution Station*

To evaluate the performance of the SFCED model for state evaluation, we compared it to the CNN and LSTM encoder-decoder models and also included two variants of SFCED: SFCED-3L (with a 3-layer fully convolutional encoder and decoder) and SFCED-5L (with a 5-layer fully convolutional encoder and decoder). For the CNN and LSTM encoder-decoders, the anomalous state scores are defined by calculating the average prediction error across all variables, and the anomalous scores for SFCED-3L and SFCED-5L are defined in the same way as for SFCED.

表 I: State Evaluation Results

Method	Pre	Rec	F1
SFCED	0.85	0.89	0.87
SFCED-3L	0.72	0.89	0.80
SFCED-5L	0.80	0.87	0.84
LSTM Encoder-Decoder	0.87	0.71	0.78
CNN	0.81	0.76	0.78

The Tab. I presents the performance metrics of various models employed for state evaluation. The SFCED model outperforms other methods with a precision of 0.85, a recall of 0.89, and an F1 score of 0.87, indicating a balanced performance in both identifying the true anomalies (high recall) and maintaining a low false positive rate (high precision). The SFCED variants, SFCED-3L and SFCED-5L, show a slight decline in performance compared to the original SFCED model, with the three-layer version (SFCED-3L) suffering more in precision, suggesting that fewer layers may not capture the complexity of the data as effectively as the standard model. The LSTM Encoder-Decoder model demonstrates a high precision but at the cost of recall, indicating that while it is accurate in the anomalies it detects, it misses a significant number of actual anomalies. This could be due to the model's potential overfitting to the normal state, leading to a less sensitive anomaly detection. The CNN model shows a balanced but overall lower performance compared to SFCED, with an F1 score matching the LSTM Encoder-Decoder model. This indicates that while CNN is capable of capturing spatial features, it may not encapsulate the temporal correlations as effectively as SFCED. Consequently, the SFCED model demonstrates a superior balance in precision and recall, validating its effectiveness for state evaluation in distribution station scenarios. It successfully identifies anomalies with minimal false positives, making it a reliable choice for real-world applications. The variants and other models offer



(a) anomaly 1 (b) anomaly 2 (c) anomaly 3 (d) anomaly 4 (e) anomaly 5

图 4: Fault Diagnosis Results

insights into the trade-offs between model complexity and detection sensitivity.

Fig. 4 clearly demonstrates the root cause diagnosis results of five anomalous events using the SFCED model. In this figure, specific rows or columns of the residual matrices correspond to reconstruction errors for different time series. For instance, Fig. 4e shows the reconstruction error of the model for each sensor data in the distribution substation. When the data from a particular sensor shows anomalies, the corresponding row-column pairs of that sensor in the figure are labelled with red boxes to indicate that these anomalies are the root cause of the fault. The accurate identification of these anomalies demonstrates the effectiveness of the model in locating the fault in the distribution substation.

Conclusion

This study introduces a novel deep neural network framework, SFCED, designed for the state evaluation and fault diagnosis in distribution stations. The framework adeptly handles multi-source heterogeneous data from distribution substation and employs signature matrices to effectively represent state information across various scales. By integrating a fully convolutional encoder-decoder structure, SFCED excels in learning deep data features, thus enhancing the accuracy of operational state assessment and timely fault detection. Comparative analysis with standard CNN and LSTM encoder-decoder models, along with SFCED variants, establishes the superior performance of SFCED, particularly in achieving a high balance between precision and recall. The research demonstrates the practical applicability of SFCED in real-life scenarios and provides new ideas for state evaluation and diagnosis of distribution substation.

参考文献

[1] F. Ahsan, N. H. Dana, S. K. Sarker, L. Li, S. Muyeen, M. F. Ali, Z. Tasneem, M. M. Hasan, S. H. Abhi, M. R. Islam *et al.*, "Data-driven next-generation smart grid towards sustainable energy evolution: techniques and technology review," *Protection and Control of Modern Power Systems*, vol. 8, no. 3, pp. 1–42, 2023.

[2] M. Khalid, "Smart grids and renewable energy systems: Perspectives and grid integration challenges," *Energy Strategy Reviews*, vol. 51, p. 101299, 2024.

[3] R. Guo, S. Meunier, C. Protopapadaki, and D. Saelens, "A review of european low-voltage distribution networks," *Renewable and Sustainable Energy Reviews*, vol. 173, p. 113056, 2023.

[4] G. Dileep, "A survey on smart grid technologies and applications," *Renewable Energy*, vol. 146, pp. 2589–2625, 2020.

[5] Z. Zhu, Y. Lei, G. Qi, Y. Chai, N. Mazur, Y. An, and X. Huang, "A review of the application of deep learning in intelligent fault diagnosis of rotating machinery," *Measurement*, vol. 206, p. 112346, 2023.

[6] I. M. Malik, A. Sharma, and R. Naayagi, "A comprehensive and practical method for transformer fault analysis with historical data trend using fuzzy logic," *IEEE Transactions on Dielectrics and Electrical Insulation*, 2023.

[7] J. Zhou, D. Nan, J. Tan, L. Zhang, Q. Zhao, and X. Ma, "State evaluation of secondary equipment based on combination weighting and cloud model," in *2020 Chinese Automation Congress (CAC)*. IEEE, 2020, pp. 6528–6532.

[8] M. Demirci, M. Saka, H. Gozde, M. Dursun, and M. C. Taplamacioglu, "Dempster shafer evidence theory application for fault diagnosis of power transformers," in *2022 9th International Conference on Electrical and Electronics Engineering (ICEEE)*. IEEE, 2022, pp. 40–44.

[9] S. Akhtar, M. Adeel, M. Iqbal, A. Namoun, A. Tufail, and K.-H. Kim, "Deep learning methods utilization in electric power systems," *Energy Reports*, vol. 10, pp. 2138–2151, 2023.

[10] H. Guo and L. Guo, "Health index for power transformer condition assessment based on operation history and test data," *Energy Reports*, vol. 8, pp. 9038–9045, 2022.

[11] N. Zhou and Y. Xu, "A multi-evidence fusion based integrated method for health assessment of medium voltage switchgears in power grid," *IEEE Transactions on Power Delivery*, vol. 38, no. 2, pp. 1406–1415, 2022.

[12] Z. Xing and Y. He, "Multimodal mutual neural network for health assessment of power transformer," *IEEE Systems Journal*, 2023.

[13] Y. Liu, H. Li, C. Zhang, and L. Zhong, "A novel framework for power transformer fault diagnosis based on kpca-smote," in *2022 9th International Conference on Condition Monitoring and Diagnosis (CMD)*. IEEE, 2022, pp. 278–283.

[14] K. Xu, S. Bao, Y. Shan, W. Ma, Q. Xu, Y. Wang, and Y. Bao, "A wavelet transform and resnet based method for sensing the operational status of secondary equipment in the distribution room," in *2023 IEEE 5th International Conference on Power, Intelligent Computing and Systems (ICPICS)*. IEEE, 2023, pp. 1–5.

[15] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[16] P. Bergmann, K. Batzner, M. Fauser, D. Sattlegger, and C. Steger, "The mvtec anomaly detection dataset: a comprehensive real-world dataset for unsupervised anomaly detection," *International Journal of Computer Vision*, vol. 129, no. 4, pp. 1038–1059, 2021.