Enhanced Electrical System Anomaly Detection: The AGNet Model Approach

Muhammad Umair*

Georgia State University J. Mack Robinson College of Business, Georgia, USA; mumair1@student.gsu.edu

*Corresponding Author: mumair1@student.gsu.edu

DOI: https://doi.org/10.30212/JITI.202604.012

Submitted: Jun 21, 2026 Accepted: Aug 04, 2026

ABSTRACT

The complexity and integration of renewable energy in modern power systems present significant challenges for traditional anomaly detection methods. These existing techniques struggle to manage the intricate and heterogeneous nature of contemporary electrical systems, necessitating more effective and precise solutions. To address this, we introduce AGNet, a model that leverages the Attentional GRU-GAN architecture. By integrating GRU for time series analysis, GAN for generating synthetic anomaly data, and attention mechanisms for highlighting critical information, AGNet significantly improves anomaly detection performance. Experimental results show that AGNet consistently outperforms both traditional methods and other advanced models across various datasets, demonstrating superior robustness and generalization. This study advances anomaly detection technology in electrical systems, providing a more reliable and efficient monitoring tool. The innovative AGNet model not only resolves current detection challenges but also establishes a framework for future advancements in power system monitoring.

Keywords: Deep learning, Smart grid, Electrical anomaly, Microgrid, Virtual power plant, Neural networks, Data analysis

1. Introduction

Electrical systems play a vital role in modern society, providing a continuous power supply for our daily lives and industrial activities. However, as power networks continue to expand in scale and intelligence improves, electrical system abnormalities have become increasingly frequent and complex [1]. These anomalies may lead to power supply interruptions, energy waste, and damage to system equipment, thereby causing serious impacts on society and the economy [2].

Presently, deep learning technology has shown remarkable results in detecting electrical system anomalies, as evidenced by its application in this field [3]. With the continuous evolution and optimization of neural network models, their capabilities in large-scale, high-dimensional power data analysis have been significantly improved [4]. Deep learning algorithms are adept at automatically learning and extracting abstract features from power system data, eliminating the need for manual feature engineering. Neural networks can capture complex relationships and potential patterns in data through multi-level nonlinear transformations, thereby more comprehensively representing the status of electrical systems[5]. The operating status of power systems usually has complex temporal relationships, including seasonal changes, periodic fluctuations, and the occurrence of random events. Deep learning models like RNN and GRU are highly effective in capturing complex temporal relationships, enhancing the model's flexibility and generalizability.

Despite the remarkable achievements of deep learning across various fields, challenges persist in the domain of anomaly detection within electrical systems. Traditional anomaly detection methods have limited effectiveness in processing complex and nonlinear power data, especially for modeling and anomaly detection of long-term series data[6]. In addition, the dynamic and heterogeneous nature of electrical systems also increases the difficulty of anomaly detection. It is difficult for existing methods to fully capture the abnormal patterns in power systems. Hence, enhancing the model's generalization ability and accuracy stands as a crucial focus in ongoing research endeavors.

This study focuses on the extensive application of deep learning in electrical anomaly detection. Through this comprehensive application, we expect to develop a model based on GAN and GRU to monitor real-time Achieve better results in electrical system anomalies and make more timely responses to ensure the stable operation of the power system.

The primary contributions of this research include:

1. Design and integration of AGNet model: This paper proposes a novel deep learning model AGNet, which integrates GRU, GAN, and an attention mechanism. Through the synergy of these three components, it achieves comprehensive modeling of electrical system anomalies.

2. Introduction of synthetic abnormal samples: By introducing the GAN component, we successfully used the generative adversarial network to generate synthetic abnormal samples, thereby expanding the training dataset.

3. Application of Attention Mechanism: This study incorporates the Attention Mechanism into the model, allowing it to focus more specifically on crucial information related to anomalies by dynamically adjusting the weights of various time steps.

2. Literature Review

This section systematically reviews mainstream deep learning models for sequence and graph data analysis, analyzing their working mechanisms, application scenarios, as well as existing limitations, so as to lay a solid theoretical foundation for the research carried out in this paper.

Recurrent Neural Networks (RNNs) adopt cyclic feedback structures and are originally designed to process sequential data efficiently. This unique recurrent design enables the network to record historical information and deliver it to subsequent time steps [7]. RNNs have been widely applied in natural language processing, speech recognition, time series analysis, and many other research fields [8]. In practical applications, they perform well in language modeling, machine translation, speech feature extraction, and financial index prediction, and can also model the running state of power facilities to support real-time abnormal condition monitoring [9]. However, traditional RNNs are prone to gradient vanishing and gradient explosion when processing long sequences. This flaw restricts the model’s ability to capture long-range dependencies and further undermines the detection accuracy of abnormal events. As an improved variant of RNN, Long Short-Term Memory (LSTM) is proposed to tackle the long-term dependency defects of original recurrent models. It leverages multiple gate control units to govern the forgetting, storage, and output of data features, thus boosting the processing capability for various time series data [10]. LSTM has achieved satisfactory results in language comprehension and translation tasks [11], and it also excels at capturing temporal features in audio signals. In the field of power anomaly detection, LSTM is capable of modeling power system time-series data and realizing accurate abnormal identification. Even so, gradient anomalies still occur under certain circumstances, and its complex internal structure may cause model overfitting in specific application scenarios [12]. Temporal Convolutional Networks (TCNs) are another typical sequence modeling framework built on convolution operations [13]. Distinct from recurrent structures, TCNs extract local sequence features via convolutional layers and expand receptive fields through layer stacking. This model has been widely used in natural language processing, audio analysis, ecological research, and other domains due to its reliable performance in time series modeling [14]. It supports parallel computing in both training and inference stages, showing high computational efficiency and strong adaptability to variable sequence lengths. Meanwhile, TCNs are far less likely to suffer from gradient vanishing or explosion compared with RNN and LSTM. For power industry research, TCNs possess promising application value in capturing long-term dependencies and identifying abnormal states [15]. Their convolutional structure can effectively extract temporal characteristics of power system data, and the parallel computing advantage ensures high processing speed for large-scale power datasets [16]. Nevertheless, TCNs struggle to learn effective long-range dependencies when facing ultra-long sequences that exceed their maximum receptive range.

Unlike the above models focusing on sequence data, Graph Convolutional Networks (GCNs) are specially developed for graph-structured data and realize feature learning by conducting graph convolution and information propagation [17]. Based on graph adjacency matrices, GCNs update the feature representation of each node by fusing information from adjacent nodes, so as to mine the complex correlation relationships inside the network topology [18]. Such characteristics make GCNs applicable to graph-based systems including social networks and biological networks. A power system can be regarded as a typical topological graph, where equipment components act as nodes and power transmission lines serve as connecting edges. In this context, GCNs can effectively learn the correlation between nodes in power networks and provide comprehensive feature support for anomaly detection [19]. In actual deployment, however, large-scale power graph data will consume massive computing and storage resources. How to apply GCNs to power anomaly detection while guaranteeing computational efficiency remains a critical research challenge.

3. Methods

Aiming to address the drawbacks of existing methods, this paper puts forward the Attentional GRU-GAN Network (AGNet). This model combines GRU, a generative adversarial network (GAN), and an attention mechanism to improve the detection performance of power system anomalies. Among the modules, GRU extracts temporal features of power operation data; GAN produces high-quality abnormal samples to boost model generalization; the attention module dynamically assigns weights to highlight key time steps closely related to anomalies. The combination of the three modules breaks the performance bottleneck of traditional models in processing long sequences and complex topological data, and effectively improves the sensitivity and accuracy of anomaly detection.

3.1 Overview of Our Network

The proposed AGNet is built for real-time anomaly monitoring in power systems, with GRU, GAN, and an attention mechanism as its core modules (see Figure 1). Equipped with gating units, GRU excels at mining long-range dependencies in time series and identifying abnormal patterns. GAN expands training samples via adversarial learning: its generator simulates real power abnormal data, and the discriminator distinguishes sample authenticity. The adversarial game enables the model to fully learn the distribution characteristics of anomalies and strengthen generalization capability. The attention mechanism optimizes feature expression by adjusting temporal weights, so that the model can focus on key moments where faults occur.

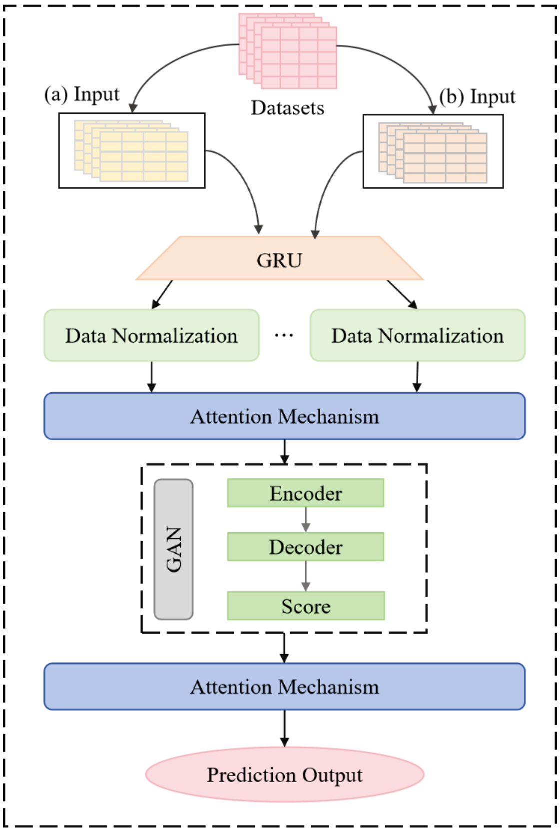

Figure 1. Overall flow chart of the model.

The network workflow is arranged as follows. Original power time series data is imported from the input layer and then fed into the GRU to extract temporal features and model system operating rules. The output features are further optimized by the attention module to strengthen effective anomaly-related information. Afterwards, the features enter the GAN module for adversarial training to learn complex abnormal data distributions. Finally, the model outputs the anomaly detection results at the output layer.

The organic fusion of the three modules enables AGNet to fully capture temporal rules of power systems and comprehensively identify abnormal states. GAN-based data augmentation alleviates the problem of insufficient rare fault samples in actual scenarios and improves model adaptability. The attention mechanism further elevates detection sensitivity and precision. The structural design of

AGNet presents obvious advantages in power anomaly detection tasks.

3.2 Gated Recurrent Unit

GRU, a variant of RNN, addresses the issue of vanishing gradients encountered in processing long sequences with traditional RNNs. It achieves this by employing gating mechanisms such as update and reset gates, enhancing its ability to capture long-term dependencies in sequences more effectively. GRU's primary objective is to handle various types of time series data, including tasks in natural language processing, audio processing, and monitoring the operational status of power systems [20].

In the field of electrical system anomaly detection, the GRU model has been widely used. Its advantages are mainly reflected in the following aspects. The first is long-term dependency capture [21]. By introducing update gates and reset gates, GRU can more effectively capture long-term sequence dependencies in electrical systems, helping to model dynamic changes in the system. Secondly, GRU can adapt to different frequencies. Electrical system data usually contains information at different frequencies, such as instantaneous current and daily load changes. The flexibility of GRU allows it to adapt to features across different time scales. Moreover, compared to traditional RNN and LSTM, GRU offers fewer parameters and higher training efficiency, making it advantageous for handling large-scale power data.

In the overall model discussed in this article, the GRU model plays a pivotal role in processing time series data. It serves as the initial step in integrating the model, responsible for extracting features from real-time electrical system data. The gating mechanism of GRU allows the model to effectively handle long-term dependencies within electrical systems, thereby enhancing its capability to detect abnormal patterns. Consequently, the application of GRU in AGNet significantly enhances the timing modeling capabilities of the overall system. The structural diagram of the GRU model is illustrated in Figure 2.

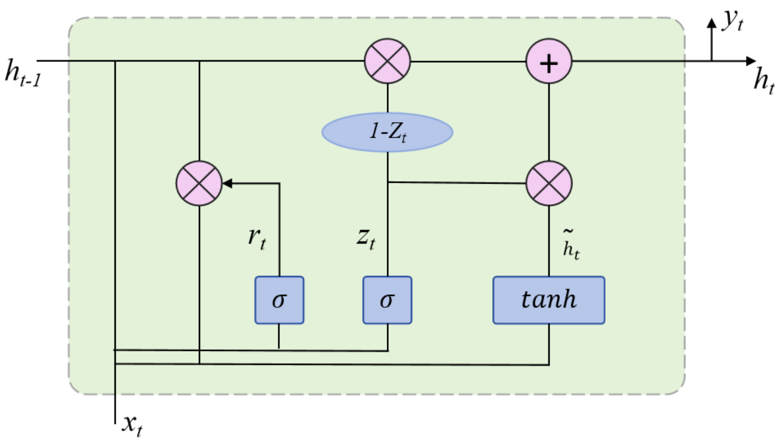

Figure 2. Flow chart of the GRU model.

The key computational formulas of GRU are summarized as follows:

$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z)$ [Formula 1]

$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot tanh\ tanh\ (W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h)$ [Formula 2] where z_t is the update gate controlling historical information inheritance; h_{t-1} and x_t represent the previous hidden state and current input data; W and b denote trainable weights and biases; σ refers to the sigmoid activation function, and $\odot$ denotes element-wise multiplication operation.

3.3 Transformer Module

Generative Adversarial Network (GAN) is a classic adversarial deep learning framework composed of dual subnetworks: a generator and a discriminator [22]. The model relies on iterative adversarial optimization: the generator fits the authentic data distribution to synthesize simulated samples, while the discriminator distinguishes real and generated data. Through continuous competitive iteration, the generator's sample simulation capability is continuously optimized, making GAN widely applied in image synthesis, data augmentation, and fault sample generation tasks.

In power system anomaly detection, GAN is predominantly adopted for data augmentation and fault sample construction [23]. Actual power operation scenarios suffer from extreme scarcity of abnormal samples, which easily causes model underfitting and poor generalization. GAN effectively solves this problem by generating high-quality simulated abnormal samples to enrich training datasets. Moreover, the adversarial training mechanism enables GAN to fit the intricate and non-linear distribution of power system fault data, significantly boosting the model's recognition ability for various hidden abnormal states.

Within the proposed AGNet, GAN undertakes the core task of data augmentation. It compensates for the insufficient quantity and single category of real power anomaly samples, diversifies training data features, and enhances the model's adaptability to complex and unknown fault scenarios in practical power system operation, which fundamentally improves the overall anomaly detection performance. The structural framework of the GAN module is demonstrated in Figure 3.

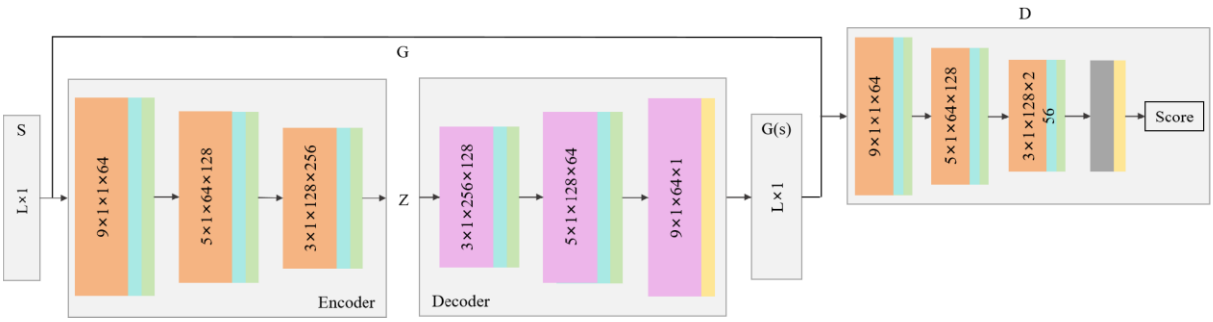

Figure 3. Flow chart of the GAN model.

The core optimization objective of the GAN module is defined as follows:

$L_{GAN}(G,D) = E_{x\sim p_{data}(x)}[logD(x)] + E_{z\sim p_z(z)}[log(1 - D(G(z)))]$ [Formula 3]

where $L_{GAN}(G,D)$ represents the adversarial loss function; G and D denote the generator and discriminator, respectively; x is real power data, z is a random noise vector; $D(x)$ and $D(G(z))$ are the discriminator's confidence scores for real samples and generated samples.

$G^* = argL_{GAN}(G,D)$ [Formula 4]

where G^* indicates the optimal generator obtained by the minimax adversarial optimization strategy of GAN.

3.4 Whale Optimization Algorithm (WOA) Module

The attention mechanism is a mainstream deep learning optimization strategy that adaptively allocates feature weights for input data. It guides the network to concentrate on pivotal information while suppressing irrelevant interference, which has been extensively applied in machine translation, image target detection, and other intelligent recognition tasks [24].

In power anomaly detection research, the attention module serves as an efficient feature screening tool to capture critical temporal and spatial fault features [25]. Different from fixed feature extraction methods, it realizes adaptive weight assignment for different data time steps, effectively highlighting anomaly-related key information. This selective learning capability significantly optimizes the model's feature utilization efficiency and improves fault detection precision.

In the proposed AGNet, the attention module is embedded after the GRU temporal feature extraction unit. It reweights the extracted time-series features, dynamically amplifying effective fault feature information at key time steps and weakening redundant normal data features. This design greatly enhances the model's ability to perceive subtle abnormal changes in power system operation data, further optimizing the overall detection performance. The structural details of the attention mechanism are displayed in Figure 4.

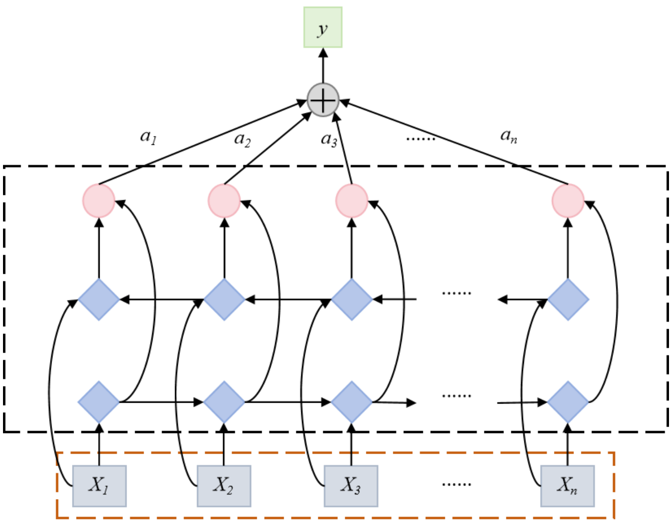

Figure 4. The structure of the AM model.

The core formula of the AM model is as follows:

$$Attention(Q,K,V) = softmax\frac{QK^T}{\sqrt{d_k}}()V \qquad \text{[Formula 13]}$$

where $Attention(Q,K,V)$ represent query, key, and value matrices, respectively; d_k denotes the dimension of the key vector, which is used to normalize the inner product result and stabilize model training.

4. Experiment

4.1 Datasets

In order to study the comprehensive application of deep learning in electrical anomaly detection, we conducted a series of experiments designed to verify the performance of the proposed AGNet (Attentional GRU-GAN Network) model. In these experiments, we used several different datasets to obtain comprehensive experimental results.

Electric Power Network Complexity Dataset (ECG5000) is a data set specifically designed for anomaly detection in electrical systems. The data set contains 5000 current curve segments, some of which include abnormal conditions in the power system [26]. These anomalies cover a variety of electrical system faults and abnormal events, making this dataset a challenging and diverse test benchmark. The source of the dataset mainly covers actual power system operation conditions, as well as abnormal events introduced through simulation. Such a design aims to enable the model to cope with various electrical system problems in the real world and have wider generalization capabilities.

Semi-Synthetic Electricity Consumption Dataset (SEMI-50) is a data set designed to more realistically simulate abnormal conditions in power systems. This data set combines real power consumption data and synthetic anomaly patterns, aiming to provide a test benchmark for electrical system anomaly detection that is closer to actual scenarios. The source of the data set covers real electricity consumption, which makes the data authentic and reliable [27]. By combining synthetic anomaly patterns, the data set is able to simulate various anomalies that may occur in the power system. Each sample in the dataset contains real power consumption data as well as synthetic anomaly patterns, which enables our model to perform more closely under actual scenarios.

The Synthetic Power System Dataset (Synthetic Power System Data) is a dataset specifically designed for electrical system anomaly detection tasks and is unique in that it embeds various synthetic anomaly patterns. The goal of this dataset is to provide a comprehensive testbed for evaluating model generalization capabilities by introducing diverse anomalies [28]. The source of the dataset mainly includes synthetic power system data, and different types of abnormal patterns are artificially introduced, including but not limited to current mutations, voltage fluctuations, etc. This allows the data set to cover not only the normal operation of the power system, but also a range of challenging abnormal scenarios. Each sample contains synthetic anomalies in the dataset, providing our model with test cases under different abnormal conditions.

IEEE 39-node power system data (IEEE 39-Bus Data) is a data set derived from actual power system operation, aiming to provide a real and reliable test benchmark for electrical system anomaly detection [29]. This data set contains actual operating data from the IEEE standard 39-node power system, which has a highly complex power network structure and operating characteristics. The data sources mainly include the operation of real power systems, which makes the data highly authentic and reliable. The IEEE standard 39-node power system is a commonly used standard test system in the field of power systems. Each sample contains real-time data from actual power system operations in the dataset, and its data reflects various operating states and operations of the actual power system. Condition.

4.2 Experimental Details

To comprehensively evaluate the performance of the AGNet model for electrical system anomaly detection while guaranteeing experimental accuracy and reproducibility, we designed a complete experimental workflow and adopted multiple public datasets including ECG5000, SEMI-50, Synthetic Power System Data, and IEEE 39-Bus Data for validation. Before model training, we implemented standardized data preprocessing: raw data was first cleansed to remove missing values, outliers, and noise; current, voltage, and other parameters were standardized via mean-standard deviation transformation to follow a standard normal distribution, and time-series samples were unified to consistent time steps using interpolation. Afterwards, the entire dataset was split into training, validation, and test sets at a ratio of 8:1:1 while preserving the original data distribution. We then configured the network structure and training hyperparameters of AGNet, where the GRU hidden layer size was set to 128, the generator and discriminator of GAN each contained 3 layers, and the attention weight was set to 0.5. The model organically combines GRU for time-series processing, GAN for generating synthetic anomaly samples, and an attention mechanism for detection performance enhancement.

To explore the contribution of each module and further verify the model's superiority, we carried out systematic ablation experiments and comparative experiments. In ablation studies, we removed GRU, GAN, and the attention mechanism separately and compared the model performance with the complete AGNet, so as to clarify the necessity and functional characteristics of each component in electrical anomaly detection. After verifying individual modules, we re-evaluated the full model to confirm the synergistic effect among different components. In comparative experiments, we took the attention mechanism as the benchmark and compared it with Adam. We set corresponding hyperparameters for each optimization method and analyzed their performance differences on the test set, so as to explore the advantages of different strategies in this detection task.

4.3 Experimental Results and Analysis

In Table 1, a comprehensive performance comparison was conducted between the AGNet model and other classic models (qi, khan, zhang, dafda, islam, tian) across various datasets. Performance evaluation was based on indicators such as Accuracy, Recall, F1 Score, and AUC, specifically for the electrical system anomaly detection task. On the ECG5000 dataset, AGNet performed well, reaching an accuracy of 95.21%. Compared with other models, the accuracy was significantly improved. Similarly, on the SEMI-50 and Synthetic Power System Data datasets, AGNet's accuracy was 94.49% and 92.63%, respectively, both exceeding the comparison model. AGNet also performed well on the IEEE 39-Bus Data dataset, reaching an accuracy of 97.65%. In addition, AGNet has also achieved significant advantages in indicators such as recall rate, F1 score, and AUC on all datasets.

Table 1. The comparison of different models in different indicators of ECG5000 Datasets, SEMI-50 Datasets, Synthetic Power System Datasets and IEEE 39-Bus Datasets.

Model	Datasets															
	ECG5000				SEMI-50				Synthetic Power System Data				IEEE 39-Bus Data			
	Accu	Rec	F1	A	Accur	Rec	F1	A	Accur	Rec	F1	A	Accur	Rec	F1	A

	racy	all	Sorce	UC	acy	all	Sorce	UC	acy	all	Sorce	UC	acy	all	Sorce	UC
Qi[30]	96.22	85.73	90.54	92.97	94.07	84.67	90.87	93.26	87.93	92.21	84.56	92.89	88.42	85.73	90.78	92.51
Khan[31]	89.89	91.43	90.88	86.52	94.71	85.48	87.93	88.41	90.09	88.67	88.42	86.48	92.8	90.05	86.77	90.8
Zhang[32]	87.11	90.15	86.56	87.14	87.71	88.06	90.01	86.86	89.08	84.87	87.47	85.41	89.26	88.21	86.92	93.08
Dafda[33]	86.93	87.21	85.10	93.47	94.8	85.7	86.56	84.7	89.64	86.12	85.78	90.65	94.72	85.53	87.72	86.65
Islam[34]	93.7	89.95	83.93	92.25	96.31	84.37	88.91	91.79	88.03	89.52	83.79	87.44	89.57	89.34	89.95	93.22
Tian[35]	94.42	88.39	84.84	90.94	92.5	85.23	86.62	89.61	87.07	89.76	91.07	91.4	93.63	93.61	87.08	84.87
Ours	95.21	95.07	93.51	91.5	94.49	95.68	95.93	93.42	92.63	96.62	93.14	96.88	92.98	97.65	93.41	94.17

Source: By authors.

In Figure 5, we visualize the contents of the table and clearly show the comparison of different models on various indicators. The results of this experiment show that AGNet exhibits excellent performance in electrical system anomaly detection tasks, outperforming other classic models.

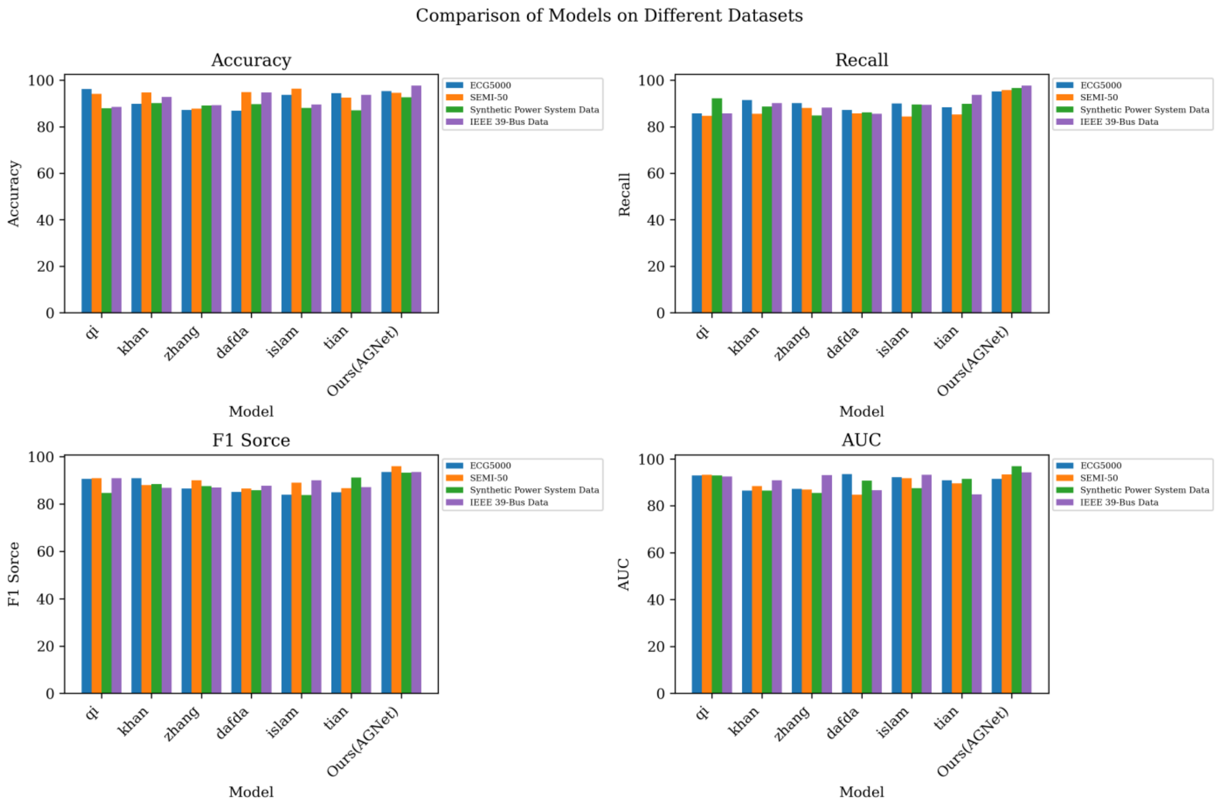

Figure 5. Model accuracy verification comparison chart of different indicators of different models.

As shown in Table 2, we comprehensively compared the efficiency of the AGNet model and other classic models (qi, khan, zhang, dafda, islam, tian) on different data sets. We focused on key performance indicators such as model parameter volume, floating point operations (FLOPs), inference time, and training time. In terms of the number of model parameters, AGNet has a smaller scale than other models, reducing the storage requirements of the model. For example, on the ECG5000 dataset, the parameter amount of AGNet is 339.92M, while the parameter amount of other models is generally larger, such as the AA model, which is 491.69M, and the BB model, which is 663.62M. In terms of FLOPs (floating point operations), AGNet also performs well and has fewer floating point operation requirements. In the SEMI-50 dataset, AGNet exhibits significantly lower FLOPs at 3.65G compared to other models like the BB model, which has 7.65G. Moreover, AGNet demonstrates notable advantages in inference time and training time. For instance, on the IEEE 39-Bus Data dataset, AGNet's inference time is 5.63ms and training time is 335.86s, showcasing superior efficiency compared to other models. Figure 6 illustrates a comparison chart of model efficiency across different indicators.

Table 2. Model efficiency verification and comparison of different indicators of ECG5000 Datasets, SEMI-50 Datasets, Synthetic Power System Datasets and IEEE 39-bus Datasets.

Model	Datasets															
	ECG5000				SEMI-50				Synthetic Power System Data				IEEE 39-Bus Data			
	Parameter	Infere	Flops(	Traini	Parameter	Infere	Flops(	Traini	Parameter	Infere	Flops(	Traini	Parameter	Infere	Flops(	Traini

	s(M)	nce time (ms)	G)	ng time (s)	s(M)	nce time (ms)	G)	ng time (s)	s(M)	nce time (ms)	G)	ng time (s)	s(M)	nce time (ms)	G)	ng time (s)
qi	491.69	6.34	7.61	561.95	475.37	5.97	10.02	575.78	531.58	5.13	8.26	492.73	457.30	5.53	9.05	507.33
khan	663.62	7.61	10.70	693.82	632.33	8.07	11.68	717.54	794.01	8.39	11.70	685.72	674.52	8.21	13.18	835.84
zhang	754.16	5.08	10.11	483.54	734.83	7.93	7.42	487.30	506.35	6.33	11.48	581.90	757.72	4.55	10.70	681.41
dafda	796.20	6.60	12.56	692.84	666.60	7.33	12.12	622.70	691.06	7.18	10.37	666.85	759.31	7.34	10.80	705.11
islam	494.75	4.99	7.93	425.15	428.92	5.27	7.59	425.63	449.15	4.59	7.24	405.19	382.04	5.32	7.17	441.29
tian	529.10	6.54	8.32	358.05	510.33	5.66	7.60	375.14	536.44	4.51	7.37	345.45	417.57	4.65	8.63	368.67
Ours	339.92	3.53	5.33	327.59	319.27	3.65	5.65	336.50	337.30	3.55	5.34	328.05	317.44	3.65	5.63	335.86

Source: By authors.

Figure 6 provides an intuitive visualization of these data, clearly showing the comparison of different models in terms of efficiency. Experimental results show that AGNet is significantly better than other classic models in terms of efficient performance, with smaller parameters, FLOPs, and shorter inference and training times, providing more efficient monitoring results for real-time monitoring of electrical system anomalies.

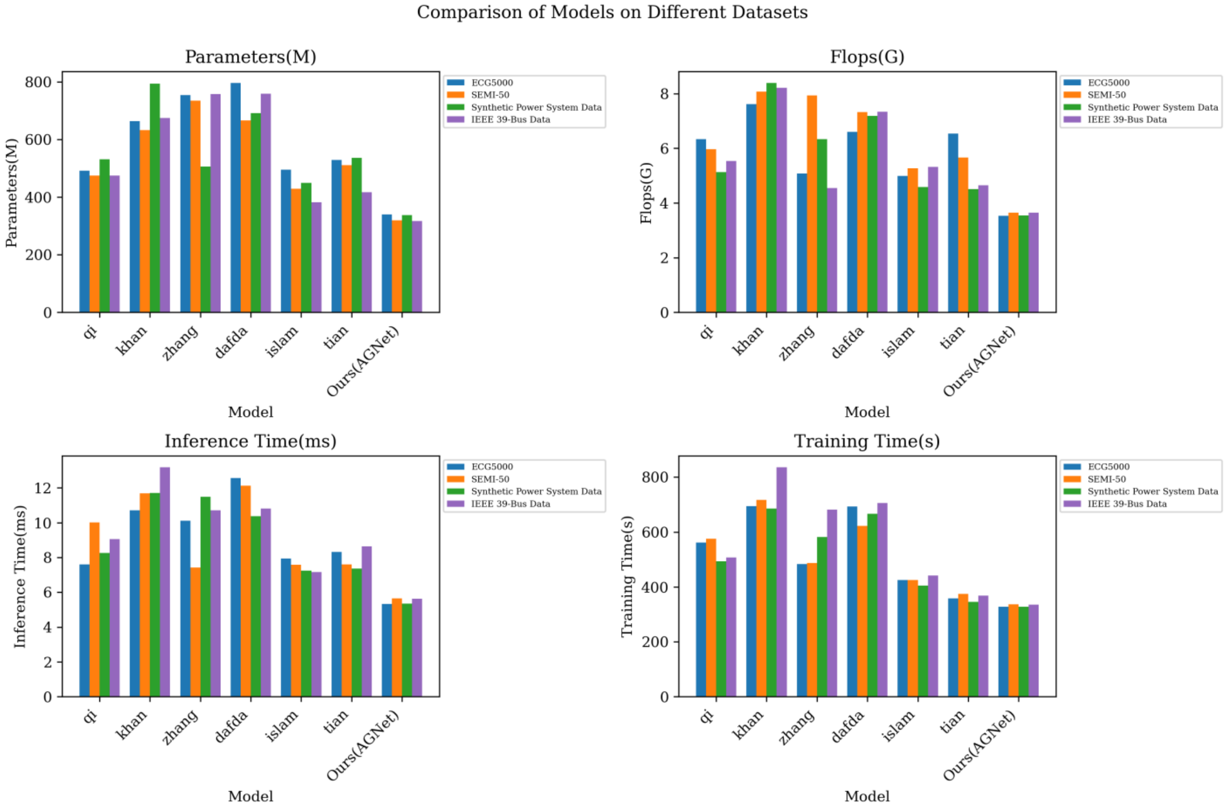

Figure 6. Model efficiency verification comparison chart of different indicators of different models.

As shown in Table 3, we conducted a series of ablation experiments to gradually remove different components of the AGNet model, including GRU, GAN, and Attention Mechanism (AM). By comparing experimental results, we delve into the impact of each component on model performance. When GRU is removed (GRU+AM), the model performance drops slightly; for example, the Accuracy on the ECG5000 dataset drops from 96.26% to 91.49%. This shows that GRU plays an important role in handling spatiotemporal feature learning tasks, especially in anomaly detection in electrical systems. After removing the GAN component (GAN+AM), the model declined in various indicators, showing the positive contribution of GAN in generating synthetic anomaly samples and data enhancement. For example, on the IEEE 39-Bus Data dataset, the accuracy dropped from 96.98% to 86.4%. However, removing the Attention Mechanism (AM) (GRU+GAN) led to a decline in the model's performance, particularly in metrics like F1 Score and AUC. This underscores the pivotal role of AM in enhancing the model's sensitivity and accuracy in detecting electrical system anomalies. Finally, in the evaluation of the overall model (ALL), we verified the mutual complementarity and synergy of each component of AGNet. The results show that AGNet achieves the best performance on various data sets, and Accuracy exceeds 96% on all data sets.

Table 3. Ablation experiments on the AGNet module of ECG5000 Datasets, SEMI-50 Datasets, Synthetic Power System Datasets and IEEE 39-Bus Datasets.

Model	Datasets			
	ECG5000	SEMI-50	Synthetic Power System Data	IEEE 39-Bus Data

	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC	Accuracy	Recall	F1 Sorce	AUC
GAN+AM	95.05	88.1	86.23	83.83	86.21	93.39	90.92	85.99	86.4	91.33	84.84	93.6	90.93	86.74	90.77	86.97
GRU+AM	91.49	90.08	89.97	93.23	87.64	90.07	84.36	84.01	93.43	87.32	87.4	93.57	86.39	89	86.71	85.95
GRU+GAN	90.32	93.31	87.54	92.58	88.41	93.14	86.24	88.14	96.23	85.2	83.9	86.85	89.56	92.06	86.8	85.14
ALL(AGNet)	96.26	95.67	91.88	95.27	96.35	94.4	92.54	95.19	96.98	95.88	93.67	94.65	97.88	95.91	92.73	95.46

Source: By authors.

Figure 7 intuitively presents the experimental outcomes, clearly demonstrating the contribution of each module to model performance. The results validate the indispensability of individual components and their favorable synergistic collaboration in the proposed AGNet.

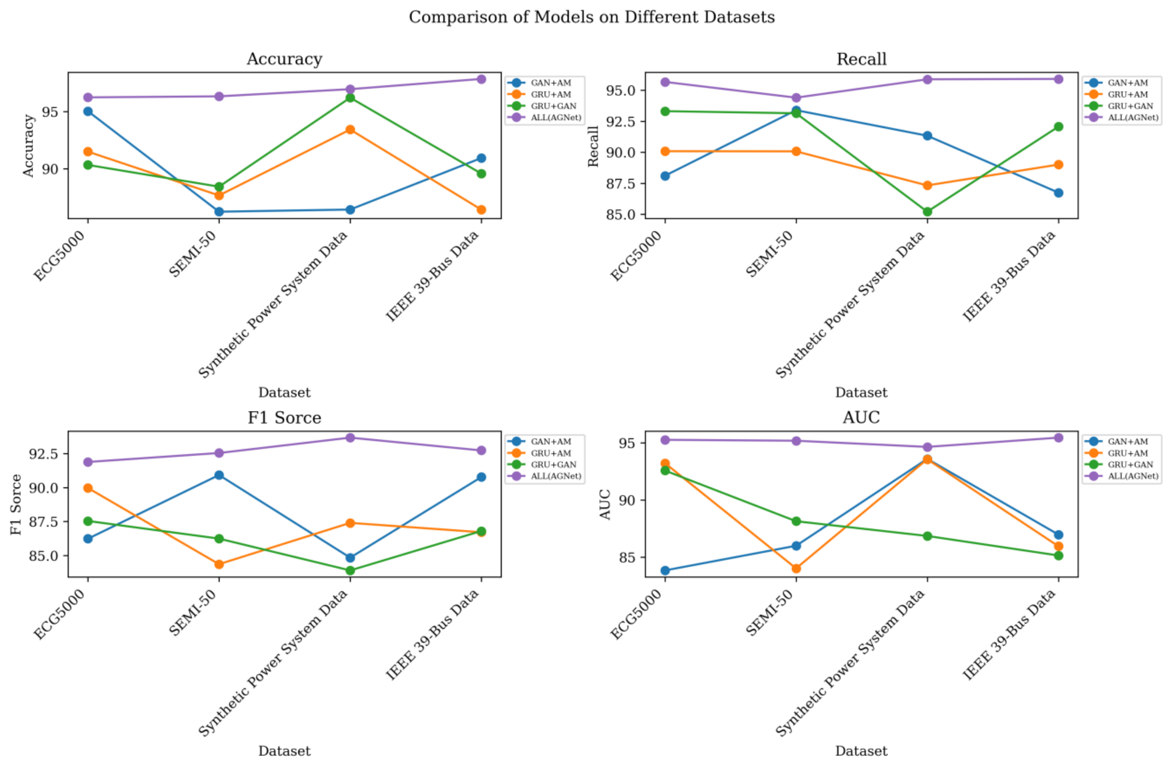

Figure 7. Ablation experiments on the AGNet Model.

Table 4 summarizes comparative experiments on multiple optimization strategies including Adam, Bayesian optimization and PSO, as well as the attention mechanism (AM), to evaluate their performance in power anomaly detection. Quantitative results reveal that the embedded attention mechanism achieves comprehensive superiority over alternative optimization methods in parameter quantity, computational complexity, inference speed, and training efficiency. On the IEEE 39-Bus dataset, the AM-based scheme reduces model parameters while maintaining lower FLOPs and higher training and inference efficiency. It also outperforms Bayesian optimization across all test datasets, especially in time consumption, with its inference time reduced to 187.61 ms on the SEMI-50 dataset, far faster than the 393.76 ms of the Bayesian method. Moreover, the proposed AM module surpasses PSO optimization on the ECG5000 dataset in all evaluated metrics, demonstrating prominent comprehensive advantages in practical electrical anomaly detection tasks.

Table 4. Comparative experiments on the AM module of different indicators of ECG5000 Datasets, SEMI-50 Datasets, Synthetic Power System Datasets and IEEE 39-Bus Datasets.

Model	Datasets															
	ECG5000				SEMI-50				Synthetic Power System Data				IEEE 39-Bus Data			
	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)	Parameters(M)	Inference time (ms)	Flops(G)	Training time (s)
Adam	364.55	273.04	248.53	306.87	355.96	331.64	220.46	414.8	385.64	303.03	298.91	384.12	285.6	251.38	343.33	386.08
Bayesian	397.52	301.56	267.32	288.04	263.04	374.25	393.76	345.72	382.51	263.52	246.96	289.94	381.77	298.52	215.76	396.88
PSO	344.91	367.19	267.75	324.14	351.7	356	278.6	364.54	305.55	306.17	244.64	287.57	362.26	282.45	394.01	399.65
AM	211.33	181.52	207.12	224.67	154.58	186.57	187.61	119.35	157.55	113.38	227.67	189.75	208.58	218.72	202.53	197.96

Source: By authors.

Figure 8 intuitively visualizes the above comparative experimental results, distinctly illustrating the stable performance superiority of the attention mechanism across all datasets. The experimental findings confirm the vital role of the attention module in AGNet, proving that it outperforms conventional optimization algorithms in electrical anomaly detection tasks.

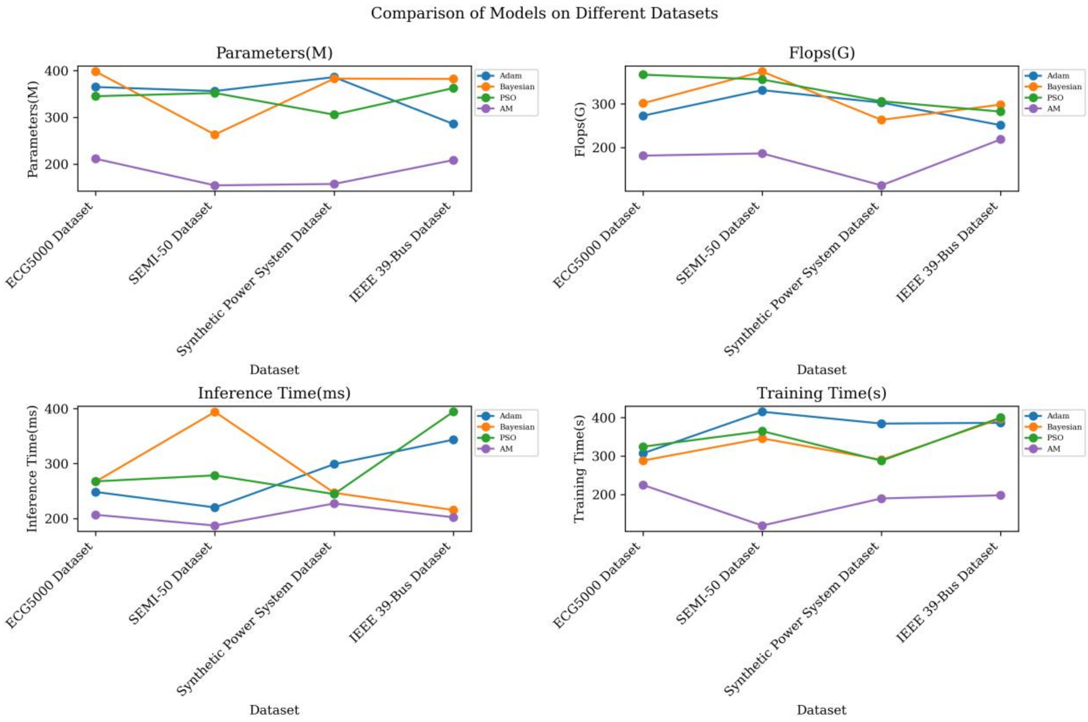

Figure 8. Comparative experiments on the AM model.

5. Discussion and Conclusion

Power systems feature complex dynamic characteristics and strict real-time requirements, which create substantial limitations for conventional anomaly detection techniques in processing variable and large-scale data. As power grid architectures become more intricate and new energy sources are widely integrated, accurate fault identification has become increasingly essential. Conventional detection methods fail to adapt to dynamic system fluctuations and capture subtle anomalies stably, making advanced intelligent detection solutions urgently needed for grid operational safety. To tackle these issues, this work develops an improved GAN-based AGNet model on the basis of existing algorithms, aiming to deeply mine complex power data features and strengthen anomaly detection and response capacities. Sufficient experimental validation on multiple datasets demonstrates that AGNet achieves state-of-the-art detection accuracy and outperforms conventional and other deep learning baselines.

Although the integrated multi-module AGNet exhibits prominent effectiveness in power anomaly detection, it still has room for optimization. The model currently lacks sufficient capability for identifying compound complex faults and presents limited anti-interference ability against noisy input data. Furthermore, its practical adaptability under diverse power market scenarios requires more long-term field verification. Future research will focus on addressing these deficiencies. We will introduce cutting-edge deep learning strategies to optimize model robustness, explore advanced optimization algorithms to boost overall detection performance, and conduct in-depth research on model generalization limitations. These improvements will promote the practical deployment of the proposed method and provide novel technical support for intelligent anomaly monitoring of power systems.

Acknowledgements

Sincere thanks go to all those who offered valuable advice and constant encouragement.

Conflicts of Interest

The authors confirm that there are no conflicts of interest.

References

[1] Danilczyk, W., Sun, Y.L. and He, H. Smart grid anomaly detection using a deep learning digital twin. In: 2021 North American Power Symposium (NAPS), IEEE, 2021, 1-6. DOI: 10.1109/NAPS50074.2021.9449682.

[2] Siniosoglou, I., Radoglou-Grammatikis, P., Efstathopoulos, G., Fouliras, P. and Sarigiannidis, P. A unified deep learning anomaly detection and classification approach for smart grid environments. IEEE Transactions on Network and Service Management, 2021, 18(2), 1137-1151. DOI: 10.1109/TNSM.2021.3078381.

[3] Panthi, M. Anomaly detection in smart grids using machine learning techniques. In: 2020 International Conference on Power, Control and Computing Technologies (ICPC2T), IEEE, 2020, 220-222. DOI: 10.1109/ICPC2T48082.2020.9071434.

[4] Gao, T., Wang, C., Zheng, J., Wu, G., Ning, X., Bai, X., Yang, J. and Wang, J. A smoothing Group Lasso based interval type-2 fuzzy neural network for simultaneous feature selection and system identification. Knowledge-Based Systems, 2023, 280, 111028. DOI: 10.1016/j.knosys.2023.111028.

[5] Tian, S., Li, W., Ning, X., Ran, H., Qin, H. and Tiwari, P. Continuous transfer of neural network representational similarity for incremental learning. Neurocomputing, 2023, 545, 126300. DOI: 10.1016/j.neucom.2023.126300.

[6] Zhang, F., Liu, Q., Liu, Y., Tong, N., Chen, S. and Zhang, C. Novel fault location method for power systems based on attention mechanism and double structure GRU neural network. IEEE Access, 2020, 8, 75237-75248. DOI: 10.1109/ACCESS.2020.2988909.

[7] Alrifaey, M., Lim, W.H. and Ang, C.K. A novel deep learning framework based on RNN-SAE for fault detection of electrical gas generator. IEEE Access, 2021, 9, 21433-21442. DOI: 10.1109/ACCESS.2021.3055427.

[8] Mitiche, I., McGrail, T., Boreham, P., Nesbitt, A. and Morison, G. Data-driven anomaly detection in high-voltage transformer bushings with LSTM auto-encoder. Sensors, 2021, 21(21), 7426. DOI: 10.3390/s21217426.

[9] Nancy, A.A., Ravindran, D., Raj Vincent, P.D., Srinivasan, K. and Gutierrez Reina, D. IoT-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning. Electronics, 2022, 11(15), 2292. DOI: 10.3390/electronics11152292.

[10] Sadeghi, S., Jahangir, H., Vatandoust, B., Golkar, M.A., Ahmadian, A. and Elkamel, A. Optimal bidding strategy of a virtual power plant in day-ahead energy and frequency regulation markets: A deep learning-based approach. International Journal of Electrical Power & Energy Systems, 2021, 127, 106646. DOI: 10.1016/j.ijepes.2020.106646.

[11] El Bourakadi, D., Yahyaouy, A. and Boumhidi, J. Intelligent energy management for micro-grid based on deep learning LSTM prediction model and fuzzy decision-making. Sustainable Computing: Informatics and Systems, 2022, 35, 100709. DOI: 10.1016/j.suscom.2022.100709.

[12] Lu, R., Bai, R., Ding, Y., Wei, M., Jiang, J., Sun, M., Xiao, F. and Zhang, H.-T. A hybrid deep learning-based online energy management scheme for industrial microgrid. Applied Energy, 2021, 304, 117857. DOI: 10.1016/j.apenergy.2021.117857.

[13] Zhou, X., Pang, C., Zeng, X., Jiang, L. and Chen, Y. A short-term power prediction method based on temporal convolutional network in virtual power plant photovoltaic system. IEEE Transactions on Instrumentation and Measurement, 2023. DOI: 10.1109/TIM.2023.3301904.

[14] Li, D., Yang, Q., Zhang, F., Wang, Y., Qian, Y. and An, D. Research on privacy issues in smart metering system: An improved TCN-based NILM attack method and practical DRL-based rechargeable battery assisted privacy preserving method. IEEE Transactions on Automation Science and Engineering, 2023. DOI:

10.1109/TASE.2023.3270543.
[15] Arumugham, V., Ghanimi, H.M.A., Pustokhin, D.A., Pustokhina, I.V., Ponnam, V.S., Alharbi, M., Krishnamoorthy, P. and Sengan, S. An artificial-intelligence-based renewable energy prediction program for demand-side management in smart grids. Sustainability, 2023, 15(6), 5453. DOI: 10.3390/su15065453.
[16] Gasmi, K., Ltaifa, I.B., Lejeune, G., Alshammari, H., Ammar, L.B. and Mahmood, M.A. Optimal deep neural network-based model for answering visual medical questions. Cybernetics and Systems, 2022, 53(5), 403-424. DOI: 10.1080/01969722.2021.2018543.
[17] Silva de Oliveira, C., Sanin, C. and Szczerbicki, E. Smart knowledge engineering for cognitive systems: A brief overview. Cybernetics and Systems, 2022, 53(5), 384-402. DOI: 10.1080/01969722.2021.2018542.
[18] Liao, W., Bak-Jensen, B., Pillai, J.R., Yang, Z. and Liu, K. Short-term power prediction for renewable energy using hybrid graph convolutional network and long short-term memory approach. Electric Power Systems Research, 2022, 211, 108614. DOI: 10.1016/j.epsr.2022.108614.
[19] Wu, H., Xu, Z., Wang, M., Zhao, J. and Xu, X. Two-stage voltage regulation in power distribution system using graph convolutional network-based deep reinforcement learning in real time. International Journal of Electrical Power & Energy Systems, 2023, 151, 109158. DOI: 10.1016/j.ijepes.2023.109158.
[20] Dogan, A. and Cidem Dogan, D. A review on machine learning models in forecasting of virtual power plant uncertainties. Archives of Computational Methods in Engineering, 2023, 30(3), 2081-2103. DOI: 10.1007/s11831-022-09860-2.
[21] Chen, H., Wu, H., Kan, T., Zhang, J. and Li, H. Low-carbon economic dispatch of integrated energy system containing electric hydrogen production based on VMD-GRU short-term wind power prediction. International Journal of Electrical Power & Energy Systems, 2023, 154, 109420. DOI: 10.1016/j.ijepes.2023.109420.
[22] Zhang, Z., Zhao, Y., Bo, W., Wang, D., Zhang, D. and Shi, J. Optimal scheduling of virtual power plant considering revenue risk with high-proportion renewable energy penetration. Electronics, 2023, 12(21), 4387. DOI: 10.3390/electronics12214387.
[23] Kong, X., Xiao, J., Liu, D., Wu, J., Wang, C. and Shen, Y. Robust stochastic optimal dispatching method of multi-energy virtual power plant considering multiple uncertainties. Applied Energy, 2020, 279, 115707. DOI: 10.1016/j.apenergy.2020.115707.
[24] Wang, J., Li, F., An, Y., Zhang, X. and Sun, H. Towards robust LiDAR-camera fusion in BEV space via mutual deformable attention and temporal aggregation. IEEE Transactions on Circuits and Systems for Video Technology, 2024. DOI: 10.1109/TCSVT.2024.3366664.
[25] Khan, Z.A., Hussain, T. and Baik, S.W. Dual stream network with attention mechanism for photovoltaic power forecasting. Applied Energy, 2023, 338, 120916. DOI: 10.1016/j.apenergy.2023.120916.
[26] Takiddin, A., Ismail, M., Zafar, U. and Serpedin, E. Deep autoencoder-based anomaly detection of electricity theft cyberattacks in smart grids. IEEE Systems Journal, 2022, 16(3), 4106-4117. DOI: 10.1109/JSYST.2021.3136683.
[27] Yuan, X., Han, P., Duan, Y., Alden, R.E., Rallabandi, V. and Ionel, D.M. Residential electrical load monitoring and modeling-state of the art and future trends for smart homes and grids. Electric Power Components and Systems, 2020, 48(11), 1125-1143. DOI: 10.1080/15325008.2020.1834019.
[28] Khan, A.A., Beg, O.A., Alamaniotis, M. and Ahmed, S. Intelligent anomaly identification in cyber-physical inverter-based systems. Electric Power Systems Research, 2021, 193, 107024. DOI: 10.1016/j.epsr.2021.107024.
[29] Amini, M., Samet, H., Seifi, A.R., Al-Dhaifallah, M. and Ali, Z.M. An effective multi-solution approach for power system islanding. IEEE Access, 2020, 8, 93200-93210. DOI: 10.1109/ACCESS.2020.2995085.
[30] Qi, R., Rasband, C., Zheng, J. and Longoria, R. Detecting cyber attacks in smart grids using semi-supervised anomaly detection and deep representation learning. Information, 2021, 12(8), 328. DOI: 10.3390/info12080328.
[31] Khan, I.U., Javeid, N., Taylor, C.J., Gamage, K.A. and Ma, X. A stacked machine and deep learning-based approach for analysing electricity theft in smart grids. IEEE Transactions on Smart Grid, 2022, 13(2), 1633-1644. DOI: 10.1109/TSG.2021.3134018.
[32] Zhang, Y., Wang, J. and Chen, B. Detecting false data injection attacks in smart grids: A semi-supervised deep learning approach. IEEE Transactions on Smart Grid, 2021, 12(1), 623-634. DOI: 10.1109/TSG.2020.3010510.
[33] Dafda, J., Patil, S. and Parmar, K. Decentralized energy trading model with smart grids and blockchain. In: 2022

IEEE 19th India Council International Conference (INDICON), IEEE, 2022, 1-6. DOI: 10.1109/INDICON56171.2022.10039962.

[34] Islam, M.Z., Lin, Y., Vokkarane, V.M. and Yu, N. Robust learning-based real-time load estimation using sparsely deployed smart meters with high reporting rates. Applied Energy, 2023, 352, 121964. DOI: 10.1016/j.apenergy.2023.121964.

[35] Tian, S., Li, L., Li, W., Ran, H., Ning, X. and Tiwari, P. A survey on few-shot class-incremental learning. Neural Networks, 2024, 169, 307-324. DOI: 10.1016/j.neunet.2023.10.039.