Edge Intelligence for Smart Grid: A Survey on Application Potentials

Hoay Beng Gooi, *Fellow*, *IEEE*, Tianjing Wang, *Member*, *IEEE*,
and Yong Tang, *Senior Member*, *IEEE*, *Fellow*, *CSEE*

***Abstract*—With the booming of artificial intelligence (AI), Internet of Things (IoT), and high-speed communication technology, integrating these technologies to innovate the smart grid (SG) further is future development direction of the power grid. Driven by this trend, billions of devices in the SG are connected to the Internet and generate a large amount of data at network edge. To reduce pressure of cloud computing and overcome defects of centralized learning, emergence of edge computing (EC) makes the computing task transfer from the network center to the network edge. When further exploring the relationship between EC and AI, edge intelligence (EI) has become one of the research hotspots. Advantages of EI in flexibly utilizing EC resources and improving AI model learning efficiency make its application in SG a good prospect. However, since only a few existing studies have applied EI to SG, this paper focuses on the application potential of EI in SG. First, the concepts, characteristics, frameworks, and key technologies of EI are investigated. Then, a comprehensive review of AI and EC applications in SG is presented. Furthermore, application potentials for EI in SG are explored, and four application scenarios of EI for SG are proposed. Finally, challenges and future directions for EI in SG are discussed. This application survey of EI on SG is carried out before EI enters the large-scale commercial stage to provide references and guidelines for developing future EI frameworks in the SG paradigm.**

***Index Terms*—Artificial intelligence, edge computing, edge intelligence, federated learning, smart grid.**

I. Introduction

DRIVEN by development and application of Internet of Things (IoT) technology, vast volumes of data are generated in smart grids (SG). To collect, transmit and process these data items, information & communication technologies (ICTs) play a vital role [1], [2]. Computing is one of the main functions of ICTs, which determines the way of data processing [3]. At present, the computing framework widely deployed in SG is cloud computing [4]. However, with increasing number of IoT devices in SG, centralized cloud computing may cause the following obstacles to development of SG: (1) Bandwidth of distribution network limits speed of cloud computing; (2) Cloud computing is challenging to meet all computing needs of millions of SG devices; and (3) Communication brought by cloud computing threatens data security.

To deal with the hidden dangers of cloud computing in continuous development of IoT, edge computing (EC) is born, which transfers computing from the network center to the network edge [5]. During implementation of SG, EC can realize a real-time collection of heterogeneous data for different devices by using container technology and providing elastic computing resources for a deep learning model. EC resource allocation can meet offline processing and analysis to ensure secure transmission and processing of various data. In addition, with the help of high-speed communication technology, EC can reduce network latency, improve utilization of network transmission bandwidth and realize efficient and stable data transmission [6]. Some devices in the SG already have functions similar to EC, such as computing, calculating, and controlling of relay protection equipment. Although these functions reflect characteristics of EC to some extent, they are not intelligent and flexible compared with actual EC [3]. In the near future, with continuous increase of distributed generations and electrical load, importance of EC in SG will become more and more prominent.

With proliferation of EC, how to combine EC with AI has become a new research hotspot, giving rise to a new concept, i.e., "edge intelligence (EI)" or "edge AI" [7], [8]. Unlike EC, which only applies AI algorithms to the edge directly, EI allocates training and inference of the AI model to different EC resources in an optimal way with consideration of economy, efficiency, and reliability. For example, federated learning can train an excellent global model through decentralized training. I. Stoica of UC Berkeley first proposed cloud edge AI is a significant development direction to achieve critical mission and personalized AI [9]. Subsequently, the concept of EI first appeared in the Gartner hype cycle in 2018 [10]. Gartner estimates EI is currently in its infancy and will peak in the next five to ten years. In recent years, several corporations, such as Google, Amazon, and Microsoft, have moved AI to the edge. To make the AI model applicable to the edge, some companies have specially designed chips for EI, such as Google edge tensor processing unit (TPU), Intel Nervana NNP, Huawei ascend 910, etc.

Rapid development trend of EI makes people see its appli-

Manuscript received April 6, 2022; revised June 21, 2022; accepted August 11, 2022. Date of online publication June 27, 2023; date of current version August 24, 2023. This work is supported by the Department of the Navy, Office of Naval Research Global under N62909-19-1-2037.

H. B. Gooi and T. J. Wang (corresponding author, email: 18810303378@163.com/tianjing.wang@ntu.edu.sg) are with Nanyang Technological University, Singapore 639798, Singapore.

Y. Tang is with China Electric Power Research Institute, Beijing 100092, China.

DOI: 10.17775/CSEEJPES.2022.02210

cation potential in the SG. Notably, individual studies have made preliminary exploration on application of EI to SG at present, from aspects of load forecasting [11], condition monitoring [12], and scenario generation [13]. However, existing studies on application of EI to SG are far from adequate. Application scenarios of EI in SG should be explicitly specified. Thus, an insightful investigation on how to achieve integration of SG and EI is warranted. This paper provides interdisciplinary values to integrate EI and SG from the prospect of SG application requirements. To clarify relevance among various concepts, a Venn diagram of EC, AI, SG, and their integrations in Fig. 1 is utilized to depict their intertwined relationships. Importance and contributions of this paper are as follows:

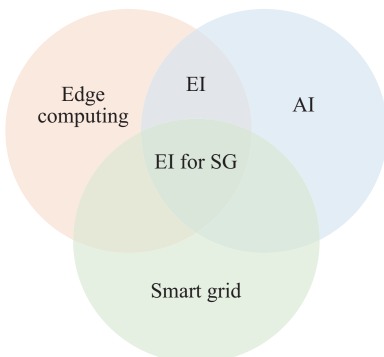

Fig. 1. Relationships of edge computing, AI, and smart grid.

1) A full technical roadmap is offered to deploy EI in SG. Specifically, concepts and characteristics of EC and EI are introduced. Moreover, frameworks and related technologies of EI in SG are presented.

2) A comprehensive review of AI and EC applications in SG is provided, which provides potential applications for EI in SG. Defects of EC applied in SG are pointed out and necessity of applying EI is illustrated.

3) Application advantages of EI in SG are analyzed, and recent advances of EI in SG are investigated. On this basis, four potential EI application scenarios in SG are proposed, i.e., decentralized-dominant application, latency-sensitive application, resource-intensive application, and security application. Enabling technologies of EI that can be leveraged in each scenario are given.

4) Challenges existing applications face in practice and future works that can be conducted are discussed from the perspectives of reliability, robustness, efficiency, sustainability, economy, and security.

The remainder of this paper is organized, as shown in Fig. 2. To clarify the relationships between EI, AI and EC, and thus illustrate application potential of EI in SG. Definition, characteristics, frameworks, and related technologies of EI are introduced in Section II. Sections III and IV provide a comprehensive review of AI and EC applied in SG. Based on associated technologies of EI and SG application scenarios, application potentials for EI in SG are given in Section V. In Section VI, challenges and future directions of EI in SG are

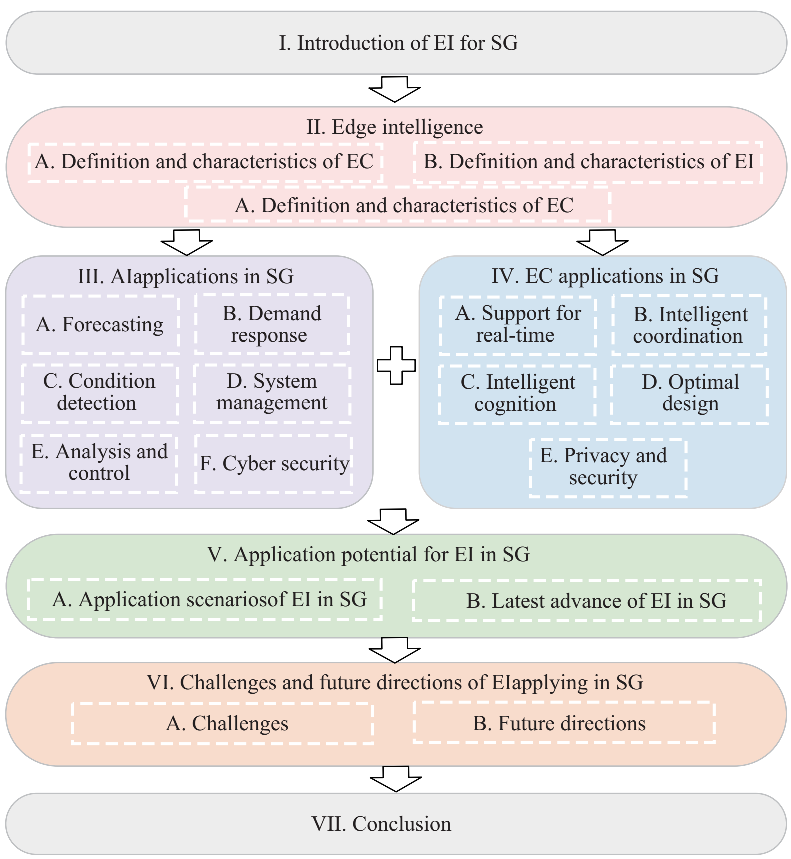

Fig. 2. Structure of this paper.

discussed. Section VII concludes this review.

II. Edge Intelligence

First, we clarify definitions and characteristics of EC and EI, and focus on differences and correlations between EI and EC. Then, based on characteristics of EI, the main framework and related technologies of model training and inference are given.

A. Definition and Characteristics of EC

1) Definition of EC

EC is a distributed computing paradigm that brings computation and data storage closer to data sources, which can avoid unnecessary communication latency and enable faster responses for end-users [14], [15]. Concepts similar to EC include cloudlets [16], fog computing [17], and mobile edge computing [18]. In addition, the concept of "edge grid" has emerged in SG, which is similar to part of the characteristics of EC, but it focuses on edge electrical devices rather than computing resources [19]. Main functions of EC include computation, communication, caching, and control. There are some devices in the power system that can realize EC-related functions, but EC is more intelligent and flexible than these devices.

2) Characteristics of EC

a) Low latency: Edge devices are located at the network edge, and the data source is very close to the transmission target, which significantly shortens communication latency. For cloud computing, communication latency is usually tens or hundreds of milliseconds, while it only takes a few milliseconds or even microsecond levels to transmit for EC [20]. Rapid development of 5G technology will further promote reduction of latency.

b) Energy efficiency: EC makes full use of existing local network resources, idle storage, and computing resources of edge devices. Moreover, data flow can be dynamically adjusted from the device to the cloud according to security policy. Data transmission volume and network bandwidth occupation can be decreased to reduce data processing cost and device energy consumption [21].

c) Privacy and security: Security of EC includes physical security and cyber security. For physical security, due to local computing characteristics of EC, impact of a single point of failure on EC is much less than that of cloud computing. For cyber security, edge servers are usually scattered in geographical locations. This is conducive to local users to manage and save data by themselves and avoid information leakage caused by data uploading and downloading [22].

d) Flexibility: In the planning stage, when the SG introduces new computing tasks, computing resources near the new tasks can be directly included in the EC, which is in line with the "plug and play" criterion of the SG. In operation phase, EC resources can be divided into multiple models for users to invoke flexibly [3].

3) Challenges of EC

In addition to the above advantages, EC inevitably has some challenges in practical application:

a) It is not easy to obtain a large number of EC resources [23]. Although edge devices have many computing resources, these computing resources generally have their inherent computing tasks. To coordinate self-computing tasks and edge computing tasks of these resources is a challenge. In addition, construction of additional EC equipment will increase investment and maintenance costs.

b) Some computing resources may not fit all computing environments. Several companies use existing computing resources for EC. However, these computing resources are often only specific to some hardware and may not be suitable for heterogeneous environments [24]. Upgrading these edge computing resources to a general computing environment brings challenges to the software platform, and it is also a significant investment.

c) Assigning tasks to near-edge nodes directly may lead to uneven allocation of computing resources and reduce computing efficiency.

B. Definition and Characteristics of EI

1) The Definition of EI

To date, there is no consensus about the official definition of EI. Some organizations and institutions define EI as running AI on an edge device [25], [26]. However, this definition is too limited. According to the relationship between AI and EC, EI can be divided into *AI for edge* and *AI on edge* [15]. *AI for edge* is optimal allocation of EC resources through AI to make EC faster and more energy-saving. *AI on edge* studies how to allocate training and inference of AI models to different EC resources in an optimal and flexible way with consideration of economy, efficiency, privacy, and reliability. This direction aims to provide a framework for training and inference of AI models on edge devices. The above definitions still cannot be used to define EI comprehensively. Considering EI is deep integration of EC and AI, EI can be defined by combining the definitions of *AI for edge* and *AI on edge*. Therefore, in this paper, EI is defined as the confluence of EC and AI that can improve operational efficiency, security, and reliability of EC or AI models by optimizing EC resource utilization. Compared with only running AI with EC, EI focuses on how to apply AI to EC in an optimal and flexible way.

2) Characteristics of EI

Apart from strengths of EC, EI has unique advantages. Compared with EC, EI can reallize not only optimal allocation of EC resources, but also enhance operational efficiency of AI model and improve utilization of EC resources through flexible deployment of AI model training and inference. Specific advantages can be described as the following:

a) Realizing optimal allocation of computing resources: EI can optimize the combination of cloud-edge-device computing resources according to application requirements and characteristics of computing resources.

b) Cognition and robustness: Since EC resources are only responsible for devices and users within their scope, they can understand specific needs and obtain nearby environmental information. In addition, by combining multiple edge models, a highly robust comprehensive model can be obtained.

c) Flexibility of model training and inference: EI separates training and inference of the AI model, so training and inference are deployed on different computing resources, making training and inference not limited by their computing resources. Besides, the model can be further segmented in training and inference to improve operational effect of the model.

d) Extensive application scenarios of AI: With EI, data associated with systems and electrical devices promote formation of AI models and enrich application scenarios of AI.

3) Challenges of EI

EI also faces some challenges in application as follows.

a) The challenge of AI for edge:

- Since the definition of EC resources in SG is not clear, a resource optimization model is not easy to build.
- Model establishment of *AI for edge* is restricted. AI-based algorithms may not work well if search space is constrained [15].
- Data related to EC resources is not easy to obtain.

b) The challenge of AI on edge:

- Raw data from variable edge devices are not readily available and may have a bias, which can affect learning performance.
- Increase in communication between models poses a threat to data security;
- How to select the scale, training frameworks and accelerator architectures poses challenges when building the model.
- How to balance optimality and efficiency challenges deploying AI algorithms on resource-constrained edge.

In a word, EI is more concerned with how to optimally deploy AI model training and inference on EC resources than EC. Based on characteristics and concerns of EI, frameworks and related technologies for model training and inference of EI are described below.

C. Frameworks and Related Technologies of EI

1) Model Training of EI

a) Frameworks: According to different training deployment modes of the AI model, the training framework can be divided into centralized, local, decentralized, and hybrid training [27], as shown in Fig. 3. Specific introductions of each framework are as follows:

- Centralized: framework deploys training of the AI model on cloud computing completely, so it is called centralized training. This mode transmits data to the cloud, avoiding resource consumption of the device and the edge, but inevitable communication cost and data security problems are very prominent. Strictly speaking, this framework does not belong to EI. However, in order to reduce computing burden of edge devices, many applications often deploy training as a centralized computing mode.

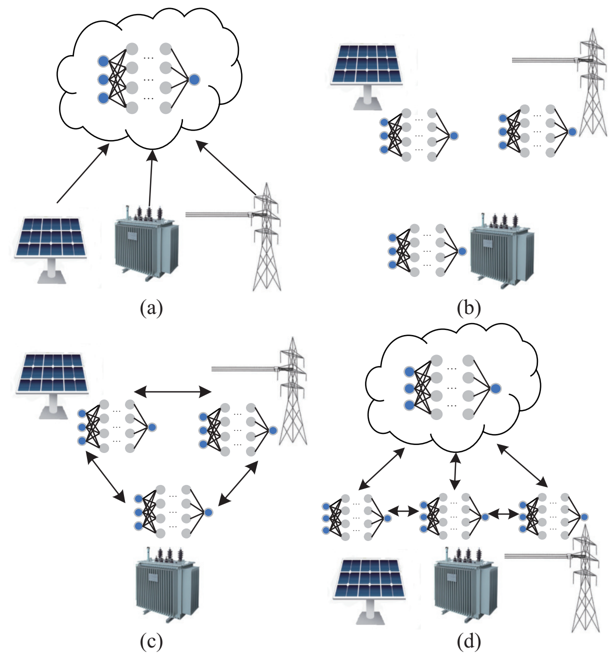

Fig. 3. Taxonomy of model training frameworks. (a) Centralized framework; (b) Local framework; (c) Decentralized framework; (d) Hybrid framework.

- Local: Each model trains on the edge or device locally, and there are few communications among models. Application of EI under this framework is a completely distributed computing mode, which is easy to lead to an overload of EC resources.
- Decentralized: It means different AI models are trained by their nearest EC resources, and then the global AI model is updated by sharing model gradient. The most typical technology in EI under this framework is federated learning, which can form a robust model while making full use of EC resources.
- Hybrid: To integrate the advantages of cloud computing and EC, centralized training and decentralized training are combined to form a hybrid training framework. Under this framework, part of the training tasks is assigned to the cloud server and the other to the edge server. This framework takes full advantage of cloud computing and EC to avoid crowding out of computing resources and allows more efficient model training for EI.

Most model training tasks in SG are deployed on the cloud. From the distributed trend of EI, the higher the degree of decentralized training, the more mature the development. Some studies have shown that edge-cloud synergy can reduce time delay and save energy consumption. However, which EI training framework is the best cannot be measured according to a standard but depends on actual needs in SG.

b) Key Technologies:

- Decentralized learning technology

A typical decentralized learning technology in EI is federated learning. Federated learning is a learning method of sharing parameters through decentralized training [28]. The main problem faced by this method is how to update parameters and whether they are updated well or not. In terms of updating gradients, a selective stochastic gradient descent protocol was proposed in [29]. However, this model does not consider parameters that may be unbalanced and non-independent distribution. To supplement the above defects, Ref. [30] proposed a FedAvg method, in which the central server will average parameters updated. For the problem of whether the parameter update is good or bad, existing methods are mainly improved from two aspects: reducing number of training cycles [31] and changing parameter update structure [32]. Gradient quantization [33] and gradient sparsification are proposed to reduce communication cost of gradient updating. Gradient quantization is a lossy compression that achieves compression by reducing accuracy of each gradient, while gradient sparsification is achieved by lowering part of the gradient. Federated learning in SG is suitable for strongly decentralized learning objects, such as distributed generation, load and individual electrical devices with monitoring equipment, which can learn prediction objects of individual edge devices and form a comprehensive prediction model containing characteristics of each edge device.

- Model splitting technology

Due to limited computing power of a single edge device, deploying a complete model to a single device can easily result in slow training. In this case, splitting the model into multiple models and deploying them on different edge devices is a solution. The role of model splitting is not only to coordinate computing resources of devices and edges but also to improve privacy. To ensure the division point of deep neural network (DNN) splitting can meet latency demand, [34] utilized the differential private mechanism and divided DNN after the first convolution layer to minimize cost of the mobile device. Moreover, a multi-split machine learning (ML) system was developed for 5G cellular networks [35]. It reformulates the complicated multi-split problem to a min-cost graph search and achieves optimal efficiency. Model splitting technology is used in SG for learning objects with complex models, high training volumes and need for privacy protection, such as condition monitoring of electrical devices and microgrid energy management.

- Training acceleration technology

One of the challenges faced by EI model training is long training time. Due to limited computing resources at the edge, training times are increasing. To mitigate this problem, transfer learning and randomized gossip algorithms can be used to shorten the training period.

Transfer learning can accelerate training and reduce computing cost of the model on the edge server. We can train a primary network first, then fine-tune it; transfer it to the target network, and train the target network. In this way, training speed of the target network can be accelerated, and energy consumption of edge devices can be reduced. [36] used federated learning to train models locally and utilized transfer learning to improve training efficiency by knowledge transfer.

To further accelerate training of decentralized training, we can leverage randomized gossip algorithms to realize a training method with fast convergence. [37] first proposed a gossip averaging algorithm, which realizes rapid convergence by exchanging information from peer to peer. Then, to make the training process faster, researchers in [38] proposed a gossip-based stochastic gradient descent algorithm. The above training acceleration techniques are also applicable to training of complex models in SG.

- EC framework advancement technology

Most existing hardware architecture, software platforms and programming frameworks are designed based on centralized computing paradigm. However, edge learning focuses on different aspects, such as energy efficiency, lightweight architecture and edge-oriented computing framework. From the perspective of hardware architecture optimization, ML processors are chips designed for edge learning tasks. [39] studied the Cortex-M microcontroller and proposed a streaming hardware accelerator to accelerate convolutional neural networks (CNN) in edge devices. [40] developed an EC platform based on FPGA and realized transfer of deep learning computing from mobile devices to edge FPGA platforms. From a software perspective, both Amazon's Greenglass and Microsoft's Azure IoT edge have implemented their software platforms or services to support EC. In terms of programming frameworks, there have been some frameworks specially designed for EC, among which MXNet [41], Tensorflow Lite, and CoreML are typical representatives. Some edge devices and edge computing systems [42] are summarized in Table I.

TABLE I
SOME EDGE DEVICES AND EDGE COMPUTING SYSTEMS

	Productions	Owners
Edge device	TPU	Google
	DianNao family	Cambrain
	Turing GPUs	NVIDIA Corporation
	7 Series FPGA	Xilinx
	HiSilicon Ascend series	Huawei
	Exynos 9820	Samsung
	Xeon D-2100	Intel
	TrueNorth	IBM
Edge computing systems	CORD	The Open Network Foundation
	EdgeX	Linux Foundation
	Akraino Edge Stack	Linux Foundation
	Azure IoT Edge	Microsoft
	AWS IoT Greengrass	Amazon
	KubeEdge	Huawei
	OpenEdge	Baidu
	OpenVDAP	Connected and Autonomous dRiving Lab
	VideoEdge	Microsoft Research

SG's edge nodes may consist of different commercial and established microprocessors. Heterogeneous computing systems provide a variety of architectural capabilities to execute an application through orchestration, with subtasks having different execution requirements [42]. One type of heterogeneous computing system is a mixed-mode machine, in which one machine can run in different parallel modes. Another type is a hybrid machine system in which a set of different kinds of high-performance machines are interconnected by high-speed links [43].

2) *Model Inference of EI*

a) Frameworks: Model inference refers to testing or online applications on new data with trained models. In consideration of cooperation among cloud, edge, and devices, the deployment framework of EI inference can be divided into the following categories [23]: edge-based mode, device-based mode, edge-device mode, and edge-cloud mode. The edge-based mode and device-based mode rely on edge and device to complete the inference. For edge-device mode, partial model is on the device, and another partial model is on the edge. Moreover, edge-cloud mode refers to partial model on the edge and another partial model on the cloud. Typically, EI deploys model inference only at the edge or the device. The latter two frameworks are complex, but model inference will be faster.

b) Key Technologies:

- Model simplification technology

To alleviate computational pressure on the edge, we can use model compression to reduce complexity of the model. At present, model compression mainly includes weight pruning, data quantification, and compact architecture design, where weight pruning [44] is the most widely used. Weight pruning is a method to reasonably remove excess weight from the trained DNN. It is to sort neurons according to the contribution of each neuron and eliminate neurons with a low contribution to realize weight pruning.

Since a high-precision DNN usually has a deep structure, it consumes tremendous computing resources to execute this model. To reduce impact of deep structure on inference speed, model early exit can take output data of the early layer as the final prediction result. A typical model early exit framework is BranchyNet [45]. Based on Branchyet, researchers have successively developed distributed DNNs [46], and Edgent [47] to navigate the accuracy-latency tradeoff, DeepIns [48] for manufacturing inspection systems.

The above model simplification techniques play an important role in online inference of complex models in SG. For example, when microgrid energy management models are applied online, inference still faces a large amount of computation due to high complexity of trained models. In this case, both weight pruning and model early exit can make the model lighter and improve speed of online operation.

- Model partition technology

To further reduce calculation pressure of DNN on a single device, the model can be divided and distributed to different devices. The key to this method is how to divide the model and where to set the partition point to obtain optimal model partition effect. For partition between the edge and device, [49]

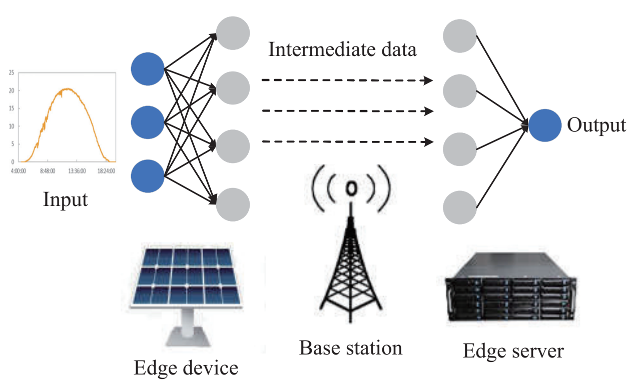

Fig. 4. Typical case for DNN partition.

partitioned DNN between device and edge to utilize nearby hybrid computing resources for real-time DNN inference. In addition, considering privacy of partition model transmission, a model partition combined with lossy feature encoding was proposed in [50]. A typical case for DNN partition between the edge and device in SG is shown in Fig. 4. PV prediction model is divided into two pieces, one running on the PV device and the other on the edge server. For partition between devices, [51] first established a microcomputing cluster based on WiFi direct technology. Then, [52] developed a partition state graph to model different partition solutions of DNNs, and proposed a neighbor effect to give the heuristic rule for the search process. When applied in SG, model partition mode can be selected according to different application scenarios and distribution of computing resources.

- Model enhance technology

Several application scenarios require fast responses when models are inferred online. However, due to complexity of the inference process, response latency is long. To achieve fast responses, edge caching is a good option. Edge caching is a method to reduce latency by buffering DNN inference results. Glimpse [53] first applied edge caching to start detection. To increase buffer capacity, [54] proposed an experience-driven edge caching method based on deep reinforcement learning, which is suitable for systems with a large number of small files. This technology can play an essential role in SG application scenarios where online response needs are high, such as demand response.

Effect of DNN inference will be affected by quality of input data. Therefore, it is crucial to filter input data. [55] proposed a method to accelerate video analysis by skipping frames with little change. This method applies a difference detector to highlight temporary differences between frames and then leverages lightweight binary classifiers to monitor differences. Besides, [56] developed an Extended Kalman Filter-based localization algorithm by edge computing to achieve a high level of accuracy and broader coverage for robot location. Filtering input data is suitable for scenarios with high data noise in SG, e.g., renewable energy generation forecasting.

- EC capacity advancement technology

To speed up AI inference, existing hardware acceleration can be used, such as CPU and GPU, as well as customized application-specific integrated circuits (ASICs) for AI, such as Google's TPU [57]. [58] proposed another customized ASIC, which focuses on efficient memory access to reduce latency and energy consumption. DNN accelerator based on a field-programmable gate array (FPGA) is another promising method because FPGA can provide fast computing while maintaining reconfigurability [59]. These customized ASIC and FPGA designs are usually more energy-saving than traditional CPU and GPU, but cost is higher.

Limitation of EC resources has led to development of technologies related to computation modes, e.g., lightweight function libraries and in-memory computing, etc. Lightweight function libraries enable more applications to run on mature operating systems in resource-limited environments of SG[60]. Some work is ongoing to build a lightweight ML library on a general-purpose chip [61]. In addition, in-memory computing is a new way of using embedded chips in the future. In-memory computing achieves high performance on a chip by storing data in RAM and processing it in parallel [62].

According to the above description, key technologies for model training and EI inference are listed in Table II.

TABLE II
KEY TECHNOLOGIES FOR MODEL TRAINING AND INFERENCE OF EI

	Types of technologies	Specific technologies
Model training	Decentralized learning technology	Federated learning
		Gradient quantization
		Gradient sparsification
	Model splitting technology	DNN splitting
	Training acceleration technology	Transfer learning
		Randomized gossip algorithms
	EC framework advancement technology	ML processors
		Heterogeneous computing
Model inference	Model simplification technology	Model compression
		Model early exit
	Model partition technology	Partition between edge and device
		Partition between devices
	Model enhance technology	Edge caching
		Input filtering
	EC capacity advancement technology	Lightweight function libraries
		In-memory computing

III. AI APPLICATIONS IN SG

With increasing penetration of renewable energy resources, SG faces an environment of flourishing uncertainty and complexity. Traditional methods have apparent defects in dealing with highly complex systems and uncertain problems. Rapid development of AI in recent years provides a new way to analyze and control SG. Deep learning (DL) can realize fitting of a highly complex system through a deep neural network to make effective classification and predictions. Moreover, reinforcement learning (RL) and deep reinforcement learning (DRL) have certain advantages in dealing with control problems with uncertainty. This section summarizes typical applications of AI in SG and discusses them from the perspective of EI.

A. Load/Renewable Energy Generation Forecasting

Predicting uncertain energy objects at different levels is crucial to reasonably arranging a power generation plan. Many studies have applied different DL to load or renewable energy generation forecasting. For load forecasting, [63] proposed a multi-scale recurrent neural network (RNN) to extract different levels of features and then learn from these features. Ref. [64] used the enhanced green Wolf optimizer to automatically architecture CNN. In addition, a few studies focus on demand forecasting at the individual building level and apply various DNNs, such as RNN-gate recurrent unit [65], unshared convolution-based deep learning [66], etc.

Renewable energy generation has characteristics of high uncertainty and intermittence, so its prediction is difficult. ML methods applied to renewable energy generation forecasting include neural networks [67]–[69], support vector machine [70], (boosted) decision tree [71], and gaussian process regression [72]. For long-term wind power forecasting, a wind power forecasting model was proposed in [73] using a tree-based learning algorithm. A robust spatio-temporal deep learning framework based on long short-term memory (LSTM) and entropy was proposed in [74] to deal with possible data pollution of photovoltaic (PV) measurement data. This method can predict PV output for multiple regions and horizons at the same time.

Models constructed in the above studies all require significant computational resources, and centralized computing would overstrain servers. However, load and new energy are naturally distributed forecasting targets, and each load or new energy field can be used as an edge computing object for EI applications. Application of EI allows for better resource utilization and less data communication for load/renewable energy generation forecasting.

B. Demand Response

Demand response is the change made by users to coordinate balance of energy supply and demand. How to predict and control energy flexibility is the key to realizing demand response. Regarding flexible load identification, [75] used neural network-integrated particle swarm optimization to identify and predict flexibility. Furthermore, researchers in [76] leveraged RNN to classify users. Demand response by controlling load is a high-dimensional control problem with randomness and partial observation, suitable for RL/DRL applications. [77] applied regular RL to resident demand response. Subsequently, a dueling deep Q network framework was proposed in [78] to optimize demand response under time of use tariff and variable electricity consumption patterns. Moreover, [79] used actor-critic to optimize demand response accounting for uncertainties in load demand. To coordinate demand response of different regions, [80] proposed a novel cooperative and decentralized reinforcement learning method, dubbed extended joint action learning, and analyzed advantages of this method compared with other centralized learning. Although the above methods are effective in achieving demand response control of loads, response time is prolonged during centralized calculations when achieving online control, which affects the control effect. For this application scenario, which is sensitive to communication time, application of EI can significantly reduce communication time of online model inference.

C. Condition Detection of Electrical Device

In SG, vast volumes of measuring devices for condition monitoring have been installed. Based on data collected by these devices, state and characteristics of the devices can be abstracted and extracted by using deep learning, and then these characteristics can be learned to realize prediction of device abnormalities, vicious attacks, fault conditions, etc. At present, ML and DL are mainly utilized to detect status of PV arrays [81], wind turbines [82], rotating electrical machines [83], transformers [84] and transmission lines [85]. An improved attention-octave convolution structure was proposed by [86] for fault detection of wind turbine converters. [87] compared a variety of ML algorithms and proposed a comprehensive detection and diagnosis system to evaluate impact of different algorithms on wind turbine fault diagnosis. Moreover, there are studies on using CNNs for deteriorated porcelain insulator detection of transmission lines [88], RNN for transformer fault diagnosis [89], etc. In the process of condition monitoring, a large amount of data is generated, which consumes a lot of resources for storage and processing. EI uses EC resources to complement cloud computing, enabling storage, processing, training and inference of condition monitoring data to be deployed at the edge, making utilization of resources more rational and efficient.

D. Microgrid Energy Management

Since the microgrid contains dynamic and variable objects such as renewable energy, varying loads, energy storage, and electric vehicles, its operation and control face significant challenges. AI algorithms have some advantages in dealing with these uncertain variables. To find optimal or approximate optimal solution for dynamic energy optimization of the microgrid, an energy management system was proposed in [90] by combining DL and an RL framework for multi-microgrid. Considering power flow equations and other operational limits, [91] presented a constrained multiagent policy learning method for optimal energy management of networked microgrids. For planning and operation of energy storage in microgrids, an RL method for controlling charge and discharge cycle of different energy storage devices was proposed in [92]. Note there are enormous amounts of control objects in the microgrid, some studies leveraged distributed control with multiagent learning to realize collaborative optimization [93], [94]. Although the above control methods are distributed, allocation of computational resources for model training and inference is still centralized. EI can allow for complete decentralization of model training and inference, making rational use of the limited resources available in the microgrid.

E. Power System Analysis and Control

Power system analysis and control are based on numerical calculation of physical models with certain assumptions and simplifications, so traditional analysis methods are difficult to predict and control accurately. Data-driven AI methods can realize model-free analysis and control through data learning. For transient stability assessment (TSA), an LSTM network was built in [95] to form a time-adaptive TSA system. Researchers in [96] transformed rotor angle into RGB pictures and constructed an online transient stability assessment system using CNNs. Besides, deep belief networks (DBNs) [97] and improved CNN-based orthogonal weight modification algorithms [98] have also been applied to TSA. In terms of prevention control, [99] proposed a deep belief network-enabled surrogate modeling for fast preventive control. [100] introduced a novel approach for transient-stability preventive control (TSPC) using graph CNN and transfer DRL. The aim is to tackle the challenges associated with the non-convergence issues observed in conventional optimization algorithms and the sluggish training pace of artificial intelligence algorithms when applied to TSPC. Taking into account the model's adaptability to diverse structures and grid configurations, [101] developed an automatic voltage control approach tailored for differential power grids. This method leverages transfer learning and DRL techniques, showcasing its ability to maintain high performance without requiring additional training, even when structural alterations occur. For emergency control, a dynamic braking and low-voltage load shedding control measure with DQN was proposed in [102]. This method can effectively maintain voltage stability and rotor angle stability of the system. To realize coordinated control of different regional power grids, [103] and [104] developed cooperative methods based on multi-agent deep reinforcement learning for load frequency control and voltage control of the multi-area power system. The power system has many controllable devices, models constructed are complex, computational resources required for training are high and real-time responsiveness of the control system is required. Application of EI to the above research allows coordination of computational resources in the power system and enhances responsiveness of the control system.

F. Cyber Security

With continuous integration of ICT in the SG, measurement and control of SG have been greatly improved. Nevertheless, it also increases risk of malicious attacks on the system, where false data injection (FDI) brings challenges to system's stable operation. Identifying and eliminating erroneous data is the key to solving this problem. Topology attacks were monitored by DRL in [105], and vulnerability of the system is analyzed. For detection of power stealing behavior, CNN was used in [106] to monitor FDI and power stealing behavior in real time. In addition, [107] utilized real-time measurement data obtained by PMUs to monitor degree of data damage through semi-supervised DL. To deal with cyber-attacks under different modes and establish prevention mechanisms, [108] proposed scenario-based two-stage spark models for cyber-attack. Moreover, [109] developed a multi-agent deep reinforcement learning algorithm for automatically discovering weak points in traditional schemes, and can identify susceptibility of most advanced detection schemes to multiple different coordinated FDI attacks on distributed communication links. AI-based approach is effective in monitoring cyber-attacks, but since model training and inference are based on centralized computing, frequent data communication leads to an increased risk of data leakage. EI can not only reduce the number of data communications, but also decrease the possibility of data leakage during model training and inference through model encryption.

IV. EC Applications in SG

According to characteristics of EC, its applications in SG include support for real-time, intelligent coordination, intelligent cognition, optimal design/resource utilization, privacy and security, and so on [110].

A. Support for Real-time

For application of supporting real-time, a distributed distribution-network fault detection method based on EC was proposed in [111], which can realize timely perception and real-time response to distribution network faults, accelerate distribution network fault processing speed, shorten outage time and improve power supply reliability and user satisfaction. Besides, [112] developed an EC-based fault location method for distribution networks with both homogenous and hybrid feeders. To overcome shortcomings of cloud computing mode in power systems, [113] proposed an EC system based on IoT SG. With help of the EC paradigm, SG based on IoT realizes connection and management of a large number of terminals, provides real-time analysis and processing of massive data, and promotes digitization of SG. For developing a low latency detection scheme of high impedance fault detection, a sustainable deep learning method with an edge device was proposed in [114], which can achieve high throughput, reduce latency and offload network traffic. In order to integrate the EC framework deeply with the infrastructure of the power system, [115] processed wide-area protection information into smaller encapsulated knowledge and transmitted it at different edge nodes. Fault incidence of each node is established and a matrix calculation is used to obtain diagnostic values for fault location, thus saving computational and communication costs.

Although EC can reduce communication latency to some extent in the above literature, computing power of edge devices is weak and the practical feasibility of applying AI models directly to EC devices is poor. Compression or splitting of models in EI technology is beneficial to solving the above problems.

B. Intelligent Coordination

In terms of intelligent coordination, [116] proposed a peer-to-peer network-based EC-based SG model in that peer-to-peer networks were applied to the EC layer, improving energy resources management efficiency and utilization of renewable energy. In addition, researchers in [117] developed a Raspberry Pi-based EC hardware prototype that coordinates agents in a low-voltage SG by solving a distributed optimal power flow.

EC is also flexible in coordination of device, edge and cloud resources. EC can adaptively optimize the way in which computing resources are composed according to requirements of the application and characteristics of computing resources. A three-layer cloud-fog computing framework of energy management was proposed in [118] for networked microgrids. Reconfiguration techniques and cloud-fog computing framework are used to quickly change topology of networked microgrids and free up line capacity to avoid feeder failures. In order to prevent the energy management system from being unable to recover in time due to a single point of failure of the microgrid, [119] developed dynamic economic scheduling based on cloud and EC framework, so the scheduling process can be carried out on the remote cloud computing platform and renewable energy inverter chip, respectively.

C. Intelligent Cognition

Compared to AI under centralized computing, EC can provide more fine-grained features for AI and achieve higher-precision cognition. EC can act as a bearer for multi-agents in AI, enabling efficient cognition of agents to realize distributed control. Microgrids contain multiple tunable power components, complicating optimization of power systems. EC-based adjustment methods can efficiently use EC resources to provide distributed intelligence solutions. [120] proposed an EC and reinforcement learning-based power control framework to control power devices, achieving goals of fast response and local autonomy. Security situational awareness is an online approach to providing security services for power system operations. [121] introduced EC between terminal and cloud to address drawbacks of centralized computing and proposes a deep reinforcement learning algorithm for EC based on multi-agent deep deterministic policy gradients to analyze security situational awareness for smart grids with minimal processing costs.

On the other hand, EC can enhance status identification capability of terminal devices. To solve the problem that existing intelligent building based on optimization participates in the control of demand response with a high cost of computing and storage, [122] proposed a cost-effective edge-cloud integration solution using reinforcement learning. Reinforcement learning uses an agent construction model that automatically learns from construction of operational data to learn optimal control strategy on the cloud infrastructure, and then distributes the strategy to edge devices for execution, which improves execution efficiency. In order to minimize resource consumption of the SG, [123] utilized EC to predict load, and proposed an intelligent resource management scheme and load forecasting model based on task unloading, improving resource allocation capacity of the SG.

Nevertheless, existing EC-related research fails to address coordination of model training and EC resources in depth. For this problem, EI provides solutions for how to deploy model training and inference on different EC resources.

D. Optimal Design/resource Utilization

With regard to optimal design/resource utilization, [124] discussed possible applications of EC in strengthening distributed optimization and control of SG, including power system asset management, distributed charging scheme and microgrid protection. Furthermore, [125] developed an EC framework to balance content generation of cameras in SG. The method could allow SG to incorporate devices that generate visual content by efficiently utilizing available resources and achieving highest Quality-of-Service.

To overcome defects of portable devices, such as large power consumption, short battery life, and intolerable delay, [126] proposed a forward central dynamic availability method. System level battery model is established by evaluating energy dissipation of IoT devices. Considering advantages of EC in energy efficiency, [127] proposed a massive multiple-input-multiple-output mobile EC framework for SG. Besides, a sequential iterative optimization algorithm was developed to jointly optimize offloading rate and transmission power to minimize energy consumption of SG.

E. Privacy and Security

Security can be divided into physical security and cyber security [3]. From the perspective of physical security, EC enhances resilience of SG in two ways [128]. On one hand, EC can support emergency communication and critical computing tasks to ensure basic system operation after a power system failure. On the other hand, after a communication system failure, the SG can use EC to transfer affected tasks to operational devices, thus avoiding cascading failures caused by large-scale communication failures. [129] explores benefits of EC for system security by using automatic voltage control as an example of how EC can guarantee security of the control process.

From the perspective of cyber security, EC enables the fundamental tasks of security monitoring and cyber security. Demand response management is a fundamental requirement for an efficient and reliable SG environment. To cope with cybersecurity issues under demand response management, [130] proposed an EC demand response management authentication scheme that achieves system resilience against various cyber-attacks. The EC framework also poses security risks to the SG. [131] developed a multi-state heterogeneous security architecture for EC-based SG to secure critical services of the SG. The system introduces intelligent scheduling, intelligent decision-making, and attack management. A dynamic scheduling algorithm based on actuator credit and heterogeneity was also proposed to improve defensive capability and security of the system.

Although EC plays a certain role in improving system security, when AI is applied to EC, data transmission of the model in the process of training and inference is still faced with the problem of data security. In this situation, EI-related technologies are required to further encrypt data to ensure data security.

V. Application Potentials for EI in SG

A. Application scenarios of EI for SG

Based on the above advantages of EI and application scenarios of AI in SG, this paper proposes four SG application scenarios that can effectively take advantage of EI, e.g., decentralized-dominant applications, latency-sensitive applications, resource-intensive applications, and security applications. Specific scenarios in SG and enabling technologies for these four application scenarios are shown in Table III.

1) Decentralized-Dominant Applications

EI is typically characterized by a decentralized computing model and can therefore be applied in decentralized-dominant scenarios. Load and renewable energy resources are naturally predictive objects in a distributed way and are typical decentralized-dominant application scenarios. Federated learning mentioned in Section II-C is leveraged to integrate models of different edges into a global model. For load forecasting, each edge node can be deployed on different types of loads, and a load forecasting model suitable for various load types can be formed by federated learning. For prediction of renewable energy generation, each node is deployed in renewable energy generating stations with different distributions. Through federated learning, a comprehensive prediction model suitable for different distributions can be obtained [132]. The global model brings together characteristics of sub models with different edges and has high robustness. Nevertheless, effect of the global model depends on edge data distribution [133]. If distribution differences of edge data are too significant, effect of the global model will be reduced. With development of the power IoT, there are more and more computing resources at power device end, which will also form a decentralized-dominant application scenario. In addition, distributed control has the advantage of decentralization because of its decentralized control mode. Training and inference of each agent in distributed control can be deployed on EC resources, forming complete decentralization. In the microgrid, distributed control is often used for control of distributed generation devices, so microgrid energy management is also a decentralized-dominant application scenario. From the perspective of hardware, because data structures of different edge devices may be heterogeneous, heterogeneous computing needs to be introduced to realize interaction of data with different structures.

2) Latency-Sensitive Applications

In view of the advantages of EI in time delay, EI can be applied to the latency-sensitive scenario of SG, including various operation control scenarios and online actions and responses in SG. Both voltage and frequency control and online demand response need low latency to ensure fast dynamic response and system performance. Taking demand response as an example, as for model inference, demand response has high requirements for communication quality. However, communication has certain uncertainties, such as noise, data loss, and latency. If model inference is implemented on the cloud, and communication quality is poor, load side cannot respond to supply side in time. Moreover, the communication path from the cloud to device is prolonged. If the model inference is deployed on the device or edge, reliability of inference will be greatly improved due to reduction of communication requirements, and latency of inference will also be reduced [27]. Due to high demand for low latency in latency-sensitive scenarios, model inference can be done using model partition and model early exit to reduce complexity of the model run by a single-edge device. To reduce latency

TABLE III
Application Scenarios and Enabling Technologies for EI in SG

Application scenarios	Specific scenarios in SG	Enabling technologies
Decentralized-dominant application	Load/renewable energy generation forecasting, condition detection of electrical device, microgrid energy management	Federated learning, input filtering, heterogeneous computing
Latency-sensitive application	Demand response, condition detection of electrical device, power system analysis and control	Model partition, model early exit, edge caching, ML processors, in-memory computing
Resource-intensive application	Condition detection of electrical device, microgrid energy management	Model splitting, transfer learning, model compression, model early exit, model partition, lightweight function libraries
Security application	Cyber security	Federated learning, model splitting

in the inference process, edge caching can be used to buffer part of the inference and achieve fast inference online. From the hardware perspective, to further improve inference speed, ML processors and in-memory computing can be leveraged to accelerate online inference.

3) Resource-Intensive Applications

In SG, billions of devices need to detect their status in real time. If their models are all trained on the cloud server during model training, it will lead to excessive computing pressure on the cloud server. Putting training on the device will also lead to excessive energy consumption due to device's capacity limitation. In this case, the hybrid mode mentioned in Section II-C can be adopted to assign some training tasks to the cloud and another part to the edge, which can alleviate computing pressure of the cloud and save energy consumption on the device [15]. However, this training framework also has some defects, e.g., increasing communication between models deployed in the cloud and the edge. Therefore, when designing the hybrid training framework, we should consider economy, reliability, and efficiency comprehensively to obtain the optimal cloud-edge-device cooperative training scheme. Resource-intensive scenarios in SG include condition detection of electrical devices and Microgrid energy management. Models built in these two scenarios are very complex, difficult to train, and resource consumption is large. To reduce complexity of model training, model splitting, transfer learning, and model compression can be used to decrease computing resource consumption of model training under a single device. For online inference, model early exit and model partition reduce complexity of model inference. In addition, calculation load can be further reduced by using lightweight function libraries.

4) Security Applications

In the inference process, failure of communication system will cause false operation of a large-scale cloud-based AI model. However, for EI, characteristics of local communication make it avoid large-scale faults and can add nearby computing resources to achieve self-healing [134]. In addition, cyber-attacks mainly come from theft and destruction during data communication. If inference of the AI model is deployed on the cloud, the cloud and device need to communicate continuously, aggravating the possibility of cyber-attacks. If the training and inference of the AI model are deployed at the edge and device, communication demand will be greatly reduced, so possibility of data being attacked will be decreased. Nevertheless, the decentralized training framework still needs a large amount of communication. In this case, data encryption methods should be introduced, such as adding noise to the data [135]. All of the above processes can be implemented through federation learning, a technique that not only enables decentralized deployment of models but also adds model and data encryption to model interactions. In addition, prevention of data leakage can also be achieved by model splitting.

B. Latest Advance of EI for SG

1) AI on Edge

Although many studies have applied AI to the edge of SG, there are few studies on how to deploy AI at the edge optimally.

Concerning load/renewable energy generation, [136] proposed a unified edge training framework combined with a light GBM algorithm for short-term prediction of PV output power. Experiments show the framework can save storage resources and energy consumption. Furthermore, a federated probabilistic forecasting scheme of solar irradiation was proposed in [132] based on deep learning, variational Bayesian inference, and federated learning, and strength of applying federated learning to uncertain prediction problems was verified. A deep learning algorithm was proposed in [137] to optimize online scheduling strategy of virtual power plants (VPPs). EC framework is designed to deal with randomness and large space characteristics of VPPs. The above research only applies AI to the EC framework or federal learning directly, without considering heterogeneous information and data leakage. Heterogeneous computing can be used to make EI applications more versatile.

When it comes to condition monitoring, a transfer learning detection method under the framework of cloud-edge cooperation was proposed in [138] to detect high-impedance faults. The edge server extracts and updates features according to synchronous measurement value provided by the phasor measurement unit. Then, all features from different distribution networks are integrated to form a unified cloud CNN model. For detection of high voltage devices, [139] proposed a data-driven framework for electrical device recognition based on infrared image. An edge-oriented generative adversarial network was developed. The edge features of electrical devices are used as a priori information to generate real infrared images. Although the above study makes full use of edge information for EI model training, it does not consider difficulty of model training due to limited EC resources. In order to shorten model training time, model splitting and model compression can be leveraged to reduce model complexity.

For detection of security operation, an abnormal target detection method combining cloud/edge fusion framework and deep learning was proposed in [140]. In view of the impact of massive heterogeneous power terminals on security situational awareness (SSA), [121] took advantage of EI to introduce edge computing between device and cloud. Furthermore, a deep reinforcement learning algorithm-based edge computing paradigm was proposed. This algorithm is utilized to analyze SSA with minimum detection error rate of SG with minimum processing cost. For power system control, [141] proposed a local power loss estimation approach for frequency emergency control. Distributed load coordination is realized by using EI and IoT technologies. This method can make distributed load provide fast and accurate frequency support for the SG. Nevertheless, the real-time requirement of frequency emergency control is high, but the above literature does not consider impact of communication delay in the process of EI online inference. Model partition and model early exit can be used to reduce complexity of model inference, and edge caching can be used to realize information buffering to ensure fast response of online inference.

Energy management within and across microgrids is complex due to many uncertainties contained in microgrids, such as renewable energy resources. [142] proposed an open framework that uses machine learning algorithms at the edge to predict energy consumption and production for energy management of smart microgrids. The framework provides good results in terms of scalability and prediction accuracy. [120] developed a power control framework combining EC and reinforcement learning, which utilized edge nodes to sense network state and control power equipment of microgrid to achieve fast response and local autonomy.

Federated learning is one of the typical techniques of EI training, as introduced in Section II-C. For application of federated learning, [143] proposed a distributed machine learning framework to detect intrusion of false data on PV DC / DC and DC / AC converters. The framework combines federated learning to support cross-device collaborative training without sharing raw data. Besides, [144] introduced federated learning into fault prediction, and the prediction model can be learned at each node through mutual communication among nodes. This strategy can achieve parallel performance and further reduce complexity of the prediction model. At present, application of federated learning in SG is in the preliminary exploration stage. Data distribution, communication cost, privacy protection, and personalized processing can also be considered in construction of federated learning.

2) AI for Edge

AI for edge utilizes AI to optimize resource allocation of EC. Usually, the resource allocation optimization model is established first, and then the AI algorithm is used to solve it. For computational tasks at multi-access edge computing (MEC) energy supply in the microgrid, [145] provided an optimization method of energy supply strategy. First, the problem is formulated as an optimization problem. Then the optimization problem is divided into two subproblems: energy-efficient tasks assignment problem and energy supply plan problem, which are solved by density-based spatial clustering and DRL, respectively. Simulation results show the performance of MEC is significantly improved under a high-precision energy supply. Subsequently, [94] leveraged a stochastic multiagent game that ensures a joint-strategy Nash equilibrium to analyze risk-aware energy scheduling problems and proved convergence of the proposed model. This approach mitigated dimensional catastrophes in state space and selected the optimal strategy among agents for risk-aware energy scheduling problems.

Table IV summarizes recent advances and EI in SG. As can be seen from the table, application of EI in SG is still at the stage of applying AI algorithms to EC only, and little research has been conducted using EI-related techniques introduced in Section II-C. Considering future needs of SG for real-time, high efficiency, and security, there is still much room for development of EI in SG research.

VI. Challenges and Future Directions of EI Applying to SG

Although EI shows excellent potential in SG and may promote rapid development of AI technology, it still faces substantial challenges in further research and implementation. According to the requirements of SG, this paper analyzes possible challenges faced by EI applying to SG from aspects of reliability, robustness, efficiency, sustainability, economy, and security, as shown in Fig. 5, and points out research directions in the future.

A. Challenges

1) Reliability

Although EI reduces data communication between device and central server, increased complexity of model training and inference reduces reliability of model operation. Therefore, reliability of model communication is a key consideration when applying EI. Unlike AI model reliability, EI-related reliability focuses on communication reliability of the model.

In terms of decentralized training, federated learning needs continuous communication between central server and each sub server to update the global model. During the communication process, if communication quality is not good enough, it may lead to incomplete model updates, affecting the model's training effect. In the process of model inference, whether for model compression or model early exit, model's integrity decreases, which may output inaccurate results during model inference. In addition, after model partition, model inference needs to be executed in different places, resulting in increased communication within the model and posing a threat to stable operation of the model.

Decentralized training composed of federated learning and the collaborative inference achieved by model partition increases communication within the model, which may reduce

TABLE IV
LATEST ADVANCE OF EI FOR SG

	Application scenarios	Specific scenarios in SG	Ref.	Year	Technologies
EI for SG	Decentralized-dominant application	PV forecasting	[97]	2021	Edge training framework + lightGBM algorithm
		PV forecasting	[103]	2020	Variational Bayesian inference + federated learning
		VPP scheduling	[104]	2020	EC framework + deep reinforcement learning
	Latency-sensitive application	Security operation detection	[107]	2020	Cloud/edge fusion framework + deep learning
		Security situational awareness	[108]	2021	Deep reinforcement learning algorithm-based EC
		Frequency emergency control	[109]	2021	Decentralized framework + cyber-physical architecture
	Resource-intensive application	High impedance faults detection	[105]	2020	Cloud-edge cooperation + transfer learning
		Electrical device recognition	[106]	2020	Edge-oriented generative adversarial network
		Microgrid energy management	[110]	2021	Machine learning algorithms at the edge
		Microgrid energy management	[111]	2021	EC framework + reinforcement learning
	Security application	False data detection on PV	[112]	2022	Federated learning + deep learning
		Fault prediction	[113]	2021	Federated learning + LSTM
	AI for Edge	Energy supply with microgrid	[114]	2019	EC + deep reinforcement learning
		Energy supply with microgrid	[115]	2021	EC + multiagent deep reinforcement learning

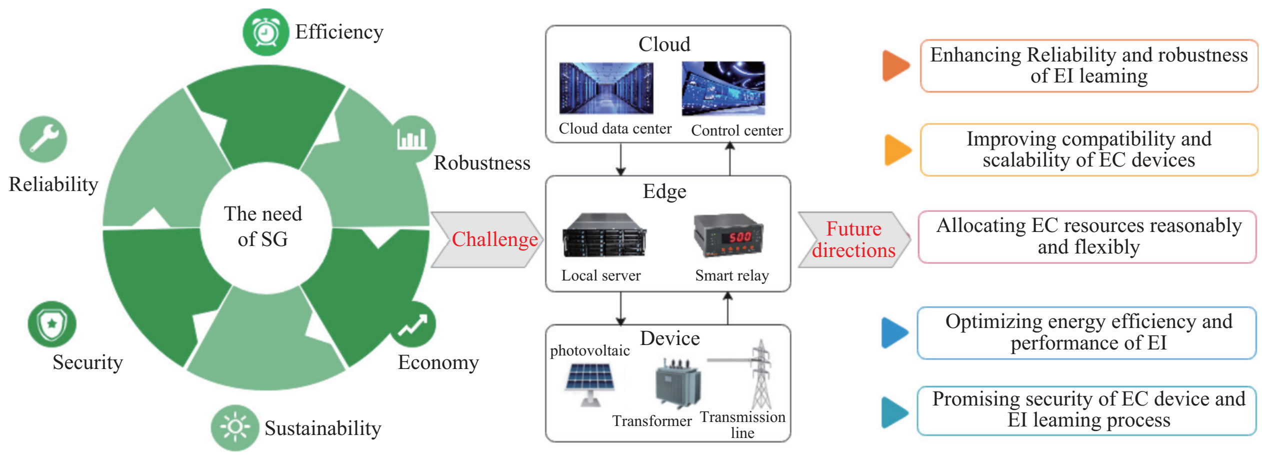

Fig. 5. Challenges and future directions of EI applying to SG.

model's reliability. Therefore, when decentralized training and model partition are applied to the AI model of SG, ensuring high reliability of internal communication of the model becomes a key consideration in the implementation process.

2) Robustness

When implementing EI, many EI-related techniques are heavily influenced by parameters and environment, so robustness is one of the challenges that EI needs to face. Compared to robustness of the AI model itself, robustness of EI needs to consider not only impact of model parameters but also influence of hardware environment.

Although federated learning can improve robustness of the model to a certain extent, application scenario of federated learning is narrow, which is only suitable for training of deep neural network models in a distributed environment. Moreover, federated learning has several requirements for number of users' devices. If the number is too small, the model's accuracy will be reduced, and training time of the model will increase. For model inference, the optimal partition point of model partition is suitable for a specific model. If the model is replaced, calculated model partition point may not be optimal. Therefore, a better training and inference mode is desired to design according to the actual application scenario to realize the high robustness of the model in training and inference.

In addition, due to a large number of heterogeneous devices using different protocols, compatibility between devices faces challenges. To improve compatibility between devices, when building the framework of EC, it can support compatibility of communication. These frameworks can be open-source and show interaction between edge devices through the interface.

3) Efficiency

Due to limitation of computing power of the edge server, if all training and inference of the models are deployed on the local server, it will inevitably lead to the problem of too slow computing speed. If the model training framework is the centralized framework, computing pressure on the central server remains unresolved. Therefore, how to rationally allocate edge computing resources to make the model run more efficiently is the challenge EI needs to face when applied. To avoid this problem, on one hand, training and inference can be split on the cloud, edge, and device to form a mutual cooperation mode. That is to say, the hybrid framework may be a preferred option. However, how to reasonably allocate computing tasks to cloud, edge, and device is another challenge. On the other hand, model compression or model early exit can be used to reduce model's complexity and improve model's inference efficiency. However, when implementing decentralized training, communication between sub-models and global models and number of sub-models will affect training efficiency. The more the sub-models, the higher the training efficiency, but the more complex the communication. In addition, computing capacity and storage capacity are two other significant limitations. Performance of the AI model and limitation of computing resources need to be weighed. Therefore, the margin of model execution complexity should be given if the model is deployed with limited computing resources.

For a complete computing task, multiple AI models may be included. Due to complexity of the model and crowded utilization of resources, it is difficult for EI to transfer these tasks. In this case, multiple AI models can be deployed on different EC servers to realize modular operation so each server can realize lightweight computing. Moreover, the modular computing method also helps build various SG tasks quickly.

4) Sustainability

In addition to sustainability of AI model, sustainability of EI mainly considers sustainability of EC resources supporting AI computing. Wide application of EI in SG requires many servers to be configured at the edge. Maintenance, renovation, and replacement of these devices is a considerable project. Due to the heterogeneity of EC resources, it is difficult to unify the maintenance scheme of devices. Besides, the scalability of EC resources is also one of the critical factors for sustainability of the system. However, while establishing an extensible EC framework, we have to provide reliable support for more devices and networks, bringing more incredible difficulty to device maintenance. Sustainability of EI is therefore an unavoidable challenge in construction of the EC framework.

Sustainability in EI is also reflected in high efficiency of devices and utilization of renewable energy resources [146]. Energy-efficient design [147], energy harvesting [148], and efficient use of renewable energy sources [149] should be a focus on to realize sustainability of EI in SG. Energy harvesting refers to looking for energy from the outside to support operation of some small devices, realizing high efficiency of devices.

5) Economy

When constructing the EI framework, economy is the key assessment indicator in system planning. Economy of EI is mainly reflected in investment of EC devices and energy consumption of EI operation. Although model training and inference on the edge/device can reduce central server's computing pressure and energy consumption, deploying multiple servers at the edge increases cost of the initial investment. In addition, model inference on edge/device increases local energy consumption and puts some pressure on local computing tasks and communication bandwidth. Therefore, it is crucial to predict local computing demand in the planning stage of configuring servers on edge. Reasonable prediction can avoid waste of initial investment. In addition, it is necessary to manage computing, storage, and communication resources of different edge devices effectively in SG. The resource management strategy can optimize comprehensive performance of the system with acceptable latency, energy consumption, and capacity.

6) Security

EI still faces challenges from both physical and cyber security perspectives. Different from security problems of AI, security of EI is reflected in security problems caused by its distributed structure and different computing resources.

Although EI enhances security of the system by reducing communication, it still faces several challenges. Geographically distributed computing resources improve the possibility of physical attacks on the system. In SG, due to the large number of edge devices, the leave one out (N-1) failure rate of all the edge devices is high. Therefore, how to reduce failure rate of distributed EC devices is a direction worthy of discussion.

In terms of cyber security, there is no unified data security protocol under the EC framework. Different edge devices may follow different protocols, which increases risk of data leakage in the process of data transmission. Therefore, how to formulate a set of unified data security protocols for EI in SG is a valuable topic.

B. Future Directions and Potential Works

According to the above possible challenges faced by applying EI to SG, future directions and potential works can be summarized as follows.

1) Enhancing Reliability and Robustness of EI Learning

- Ensure high reliability of internal communications of the model when decentralized training and model partition are applied.
- Find a better training and inference mode according to the specific application scenario to realize high robustness.

2) Improving Compatibility and Scalability of EC Devices

- Improve compatibility between devices when building the framework of EC and EI.
- Enhance scalability of EC resources.

3) Allocating EC Resources Reasonably and Flexibly

- Allocate computing tasks to cloud, edge, and device reasonably.
- Give the margin of model execution complexity when the model is deployed with limited computing resources.
- Deploy multiple AI models on different EC servers to realize modular operation.

4) Optimizing Energy Efficiency and Performance of EI

- Develop energy-efficient design, energy harvesting, and efficient use of renewable energy sources.
- Predict local computing demands in the planning stage of configuring servers on edge.
- Optimize comprehensive performance of the system with acceptable latency, energy consumption, and capacity.

5) Promising Security of EC Device and EI Learning Process

- Reduce failure rate of distributed EC devices.
- Formulate a set of unified data security protocols for EI in SG.

VII. Conclusion

Driven by boosting of AI and IoT, EC, which can transfer computing tasks from cloud to edge, is rising rapidly. The combination of EC and AI promotes the birth of EI. With a growing body of IoT devices connected to the SG, EI has great application potential in the SG. In this survey, we aim to discuss application advances and potentials for EI in SG. First, we give an in-depth insight into EI in SG, including definition, characteristics, frameworks, and related technologies. Next, we summarized a comprehensive review of AI and EC in SG. Further, we propose application advantages of EI in four scenarios, which provides a reference for further research.

Since EI technology is still in the initial development stage and has few existing applications in SG, it faces many challenges in the next exploration and practical application. According to needs of future development of SG, we clarify challenges faced by application of EI in SG from the aspects of reliability, robustness, efficiency, sustainability, economy, and security, and point out improvements and research directions that can be made in the future.

References

[1] V. C. Gungor, D. Sahin, T. Kocak, S. Ergut, C. Buccella, C. Cecati, and G. P. Hancke, "Smart grid technologies: communication technologies and standards," *IEEE Transactions on Industrial Informatics*, vol. 7, no. 4, pp. 529–539, Nov. 2011.

[2] G. I. Shokiraliyevich, "Role of information and communication technologies in accounting and digital economy," *South Asian Journal of Marketing & Management Research*, vol. 11, no. 5, pp. 17–20, Jun. 2021.

[3] C. Feng, Y. Wang, Q. X. Chen, Y. Ding, G. Strbac, and C. Q. Kang, "Smart grid encounters edge computing: opportunities and applications," *Advances in Applied Energy*, vol. 1, pp. 100006, Feb. 2021.

[4] A. Mohammed and D. Syed, "Cloud computing for smart grid," in *Smart Grid and Enabling Technologies*, S. S. Refaat, O. Ellabban, S. Bayhan, H. Abu-Rub, F. Blaabjerg, and M. M. Begovic, Eds. Hoboken: John Wiley & Sons Ltd., 2021, pp. 333–357.

[5] W. S. Shi, J. Cao, Q. Zhang, Y. H. Z. Li, and L. Y. Xu, "Edge computing: vision and challenges," *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, Oct. 2016.

[6] T. Zhang, Y. K. Li, and C. L. P. Chen, "Edge computing and its role in industrial internet: methodologies, applications, and future directions," *Information Sciences*, vol. 557, pp. 34–65, May 2021.

[7] X. F. Wang, Y. W. Han, C. Y. Wang, Q. Y. Zhao, X. Chen, and M. Chen, "In-edge AI: intelligentizing mobile edge computing, caching and communication by federated learning," *IEEE Network*, vol. 33, no. 5, pp. 156–165, Sep./Oct. 2019.

[8] Ding A Y, Peltonen E, Meuser T, et al., "Roadmap for edge ai: A dagstuhl perspective," in *Proceedings of ACM SIGCOMM Computer Communication Review*, vol. 52, no.1, 2022, pp. 28–33.

[9] I. Stoica, D. Song, R. A. Popa, D. Patterson, M. W. Mahoney, R. Katz, A. D. Joseph, M. Jordan, J. M. Hellerstein, J. E. Gonzalez, K. Goldberg, A. Ghodsi, D. Culler, and P. Abbeel. (2017, Dec.). A Berkeley view of systems challenges for AI. [Online]. Available: https://arxiv.org/abs/1712.05855.

[10] Gartner. (2019, Nov.). 5 Trends emerge in the Gartner hype cycle for emerging technologies. [Online]. Available: https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/.

[11] Y. L. Tun, K. Thar, C. M. Thwal, and C. S. Hong, "Federated learning based energy demand prediction with clustered aggregation," in *Proceedings of 2021 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 2021, pp. 164–167.

[12] G. C. Montanari, R. Hebner, P. Seri and R. Ghosh, "Self-Assessment of Health Conditions of Electrical Assets and Grid Components: A Contribution to Smart Grids," in *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1206–1214, March 2021.

[13] Y. Li, J. Z. Li, and Y. Wang, "Privacy-preserving spatiotemporal scenario generation of renewable energies: a federated deep generative learning approach," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2310–2320, Apr. 2022.

[14] J. S. Chen and X. K. Ran, "Deep learning with edge computing: a review," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1655–1674, Aug. 2019.

[15] S. G. Deng, H. L. Zhao, W. J. Fang, J. W. Yin, S. Dustdar, and A. Y. Zomaya, "Edge intelligence: the confluence of edge computing and artificial intelligence," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7457–7469, Aug. 2020.

[16] L. Zhao, W. Sun, Y. P. Shi, and J. J. Liu, "Optimal placement of cloudlets for access delay minimization in SDN-based internet of things networks," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1334–1344, Apr. 2018.

[17] M. Mukherjee, L. Shu, and D. Wang, "Survey of fog computing: fundamental, network applications, and research challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 3, pp. 1826–1857, thirdquarter, 2018.

[18] X. Z. Lai, L. S. Fan, X. F. Lei, Y. S. Deng, G. K. Karagiannidis, and A. Nallanathan, "Secure mobile edge computing networks in the presence of multiple eavesdroppers," *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 500–513, Jan. 2022.

[19] D. Divan, R. Moghe, and A. Prasai, "Power electronics at the grid edge: the key to unlocking value from the smart grid," *IEEE Power Electronics Magazine*, vol. 1, no. 4, pp. 16–22, Dec. 2014.

[20] W. B. Dai, H. Nishi, V. Vyatkin, V. Huang, Y. Shi, and X. P. Guan, "Industrial edge computing: enabling embedded intelligence," *IEEE Industrial Electronics Magazine*, vol. 13, no. 4, pp. 48–56, Dec. 2019.

[21] C. F. Jiang, T. T. Fan, H. H. Gao, W. S. Shi, L. K. Liu, C. Cérin, and J. Wan, "Energy aware edge computing: a survey," *Computer Communications*, vol. 151, pp. 556–580, Feb. 2020.

[22] A. Alwarafy, K. A. Al-Thelaya, M. Abdallah, J. Schneider, and M. Hamdi, "A survey on security and privacy issues in edge-computing-assisted internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4004–4022, Mar. 2021.

[23] R. C. Xie, Q. Q. Tang, S. Qiao, H. Zhu, F. R. Yu, and T. Huang, "When serverless computing meets edge computing: architecture, challenges, and open issues," *IEEE Wireless Communications*, vol. 28, no. 5, pp. 126–133, Oct. 2021.

[24] S. S. Liu, L. K. Liu, J. Tang, B. Yu, Y. F. Wang, and W. S. Shi, "Edge computing for autonomous driving: opportunities and challenges," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1697–1716, Aug. 2019.

[25] IEC white paper edge intelligence. [Online]. Available: https://www.iec.ch/system/files/2019--09/content/media/files/iec_wp_edge_intelligence_en_lr.pdf.

[26] S. Lin, Z. Zhou, Z. F. Zhang, X. Chen, and J. S. Zhang, *Edge Intelligence in the Making: Optimization, Deep Learning, and Applications*, Cham: Springer, 2021.

[27] Z. Zhou, X. Chen, E. Li, L. K. Zeng, K. Luo, and J. S. Zhang, "Edge intelligence: paving the last mile of artificial intelligence with edge computing," *Proceedings of the IEEE*, vol. 107, no. 8, pp. 1738–1762, Aug. 2019.

[28] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, May 2020.

[29] R. Shokri and V. Shmatikov, "Privacy-preserving deep learning," in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 2015, pp. 1310–1321.

[30] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 2017, pp. 1273–1282.

[31] L. U. Khan, S. R. Pandey, N. H. Tran, W. Saad, Z. Han, M. N. H. Nguyen, and C. S. Hong, "Federated learning for edge networks: resource optimization and incentive mechanism," *IEEE Communications Magazine*, vol. 58, no. 10, pp. 88–93, Oct. 2020.

[32] W. Liu, L. Chen, Y. F. Chen, and W. Y. Zhang, "Accelerating federated learning via momentum gradient descent," *IEEE Transactions on Parallel and Distributed Systems*, vol. 31, no. 8, pp. 1754–1766, Aug. 2020.

[33] Y. Q. Du, S. Yang, and K. B. Huang, "High-dimensional stochastic gradient quantization for communication-efficient edge learning," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2128–2142, Mar. 2020.

[34] Y. L. Mao, S. H. Yi, Q. Li, J. H. Feng, F. Y. Xu, and S. Zhong, "A privacy-preserving deep learning approach for face recognition with edge computing," in *Proceedings of USENIX Workshop on Hot Topics in Edge Computing*, 2018, pp. 1–6.

[35] S. Wang, X. Y. Zhang, H. Uchiyama, and H. Matsuda, "HiveMind: towards cellular native machine learning model splitting," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 2, pp. 626–640, Feb. 2022.

[36] Y. Y. Cheng, J. Y. Lu, D. Niyato, B. Lyu, J. W. Kang, and S. M. Zhu, "Federated transfer learning with client selection for intrusion detection in mobile edge computing," *IEEE Communications Letters*, vol. 26, no. 3, pp. 552–556, Mar. 2022.

[37] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.

[38] A. Koloskova, S. Stich, and M. Jaggi, "Decentralized stochastic optimization and gossip algorithms with compressed communication," in *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 3478–3487.

[39] L. Du, Y. Du, Y. L. Li, J. J. Su, Y. C. Kuan, C. C. Liu, and M. C. F. Chang, "A reconfigurable streaming deep convolutional neural network accelerator for internet of things," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 1, pp. 198–208, Jan. 2018.

[40] T. Belabed, M. G. F. Coutinho, M. A. C. Fernandes, C. V. Sakuyama, and C. Souani, "User driven FPGA-based design automated framework of deep neural networks for low-power low-cost edge computing," *IEEE Access*, vol. 9, pp. 89162–89180, Jun. 2021.

[41] H. J. Yang H, M. Fritzsche, C. Bartz, and C. Meinel, "BMXNet: an open-source binary neural network implementation based on MXNet," in *Proceedings of the 25th ACM International conference on Multimedia*, 2017, pp. 1209–1212.

[42] Z. Q. Chang, S. B. Liu, X. X. Xiong, Z. H. Cai, and G. Q. Tu, "A survey of recent advances in edge-computing-powered artificial intelligence of things," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 13849–13875, Sep. 2021.

[43] D. Wang, B. Bai, K. Lei, W. B. Zhao, Y. P. Yang, and Z. Han, "Enhancing information security via physical layer approaches in heterogeneous IoT with multiple access mobile edge computing in smart city," *IEEE Access*, vol. 7, pp. 54508–54521, May 2019.

[44] X. L. Ma, F. M. Guo, W. Niu, X. Lin, J. Tang, K. S. Ma, B. Ren, and Y. Z. Wang, "PCONV: the missing but desirable sparsity in DNN weight pruning for real-time execution on mobile devices," in *Proceedings of the 34th AAAI Conference on Artificial Intelligence*, 2020, pp. 5117–5124.

[45] R. G. Pacheco and R. S. Couto, "Inference time optimization using BranchyNet partitioning," in *Proceedings of 2020 IEEE Symposium on Computers and Communications (ISCC)*, 2020, pp. 1–6.

[46] C. H. Chiang, P. F. Liu, D. W. Wang, D. Y. Hong, and J. J. Wu, "Optimal branch location for cost-effective inference on branchynet," in *Proceedings of 2021 IEEE International Conference on Big Data (Big Data)*, 2021, pp. 5071–5080.

[47] E. Li, Z. Zhou, and X. Chen, "Edge intelligence: on-demand deep learning model co-inference with device-edge synergy," in *Proceedings of the 2018 Workshop on Mobile Edge Communications*, 2018, pp. 31–36.

[48] L. Z. Li, K. Ota, and M. X. Dong, "Deep learning for smart industry: efficient manufacture inspection system with fog computing," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 10, pp. 4665–4673, Oct. 2018.

[49] Y. P. Kang, J. Hauswald, C. Gao, A. Rovinski, T. Mudge, J. Mars, and L. J. Tang, "Neurosurgeon: collaborative intelligence between the cloud and mobile edge," *ACM SIGPLAN Notices*, vol. 52, no. 4, pp. 615–629, Apr. 2017.

[50] J. H. Ko, T. Na, M. F. Amir, and S. Mukhopadhyay, "Edge-host partitioning of deep neural networks with feature space encoding for resource-constrained internet-of-things platforms," in *Proceedings of the 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2018, pp. 1–6.

[51] J. C. Mao, Z. D. Yang, W. Wen, C. P. Wu, L. H. Song, K. W. Nixon, X. Chen, H. Li, and Y. R. Chen, "MeDNN: a distributed mobile system with enhanced partition and deployment for large-scale DNNs," in *Proceedings of 2017 IEEE/ACM International Conference on Computer-Aided Design*, 2017, pp. 751–756.

[52] H. L. Wang, B. Guo, J. Q. Liu, S. C. Liu, Y. G. Wu, and Z. W. Yu, "Context-aware adaptive surgery: a fast and effective framework for adaptative model partition," in *Proceedings of ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2021, pp. 131.

[53] T. Y. H. Chen, L. Ravindranath, S. Deng, P. Bahl, and H. Balakrishnan, "Glimpse: continuous, real-time object recognition on mobile devices," in *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, 2015, pp. 155–168.

[54] C. H. Song, W. X. Xu, T. T. Wu, S. M. Yu, P. Zeng, and N. Zhang, "QoE-driven edge caching in vehicle networks based on deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 5286–5295, Jun. 2021.

[55] D. Kang, J. Emmons, F. Abuzaid, P. Bailis, and M. Zaharia, "NoScope: optimizing neural network queries over video at scale," *Proceedings of the VLDB Endowment*, vol. 10, no. 11, pp. 1586–1597, Aug. 2017.

[56] I. Ullah, S. Y. Qian, Z. X. Deng, and J. H. Lee, "Extended Kalman filter-based localization algorithm by edge computing in wireless sensor networks," *Digital Communications and Networks*, vol. 7, no. 2, pp. 187–195, May 2021.

[57] N. P. Jouppi, C. Young, N. Patil, D. Patterson, G. Agrawal, R. Bajwa, S. Bates, S. Bhatia, N. Boden, A. Borchers, R. Boyle, P. L. Cantin, C. Chao, C. Clark, J. Coriell, M. Daley, M. Dau, J. Dean, B. Gelb, T. V. Ghaemmaghami, R. Gottipati, W. Gulland, R. Hagmann, C. R. Ho, D. Hogberg, J. Hu, R. Hundt, D. Hurt, J. Ibarz, A. Jaffey, A. Jaworski, A. Kaplan, H. Khaitan, D. Killebrew, A. Koch, N. Kumar, S. Lacy, J. Laudon, J. Law, D. Le, C. Leary, Z. Y. Liu, K. Lucke, A. Lundin, G. MacKean, A. Maggiore, M. Mahony, K. Miller, R. Nagarajan, R. Narayanaswami, R. Ni, K. Nix, T. Norrie, M. Omernick, N. Penukonda, A. Phelps, J. Ross, M. Ross, A. Salek, E. Samadiani, C. Severn, G. Sizikov, M. Snelham, J. Souter, D. Steinberg, A. Swing, M. Tan, G. Thorson, B. Tian, H. Toma, E. Tuttle, V. Vasudevan, R. Walter, W. Wang, E. Wilcox, and D. H. Yoon, "In-datacenter performance analysis of a tensor processing unit," in *Proceedings of the 44th Annual International Symposium on Computer Architecture*, 2017, pp. 1–12.

[58] Z. D. Du, R. Fasthuber, T. S. Chen, P. Ienne, L. Li, T. Luo, X. B. Feng, Y. J. Chen, and O. Temam, "ShiDianNao: shifting vision processing closer to the sensor," *ACM SIGARCH Computer Architecture News*, vol. 43, no. 3, pp. 92–104, Jun. 2015.

[59] S. Q. Wang, L. Wang, Y. Deng, Z. J. Yang, S. S. Guo, Z. Y. Kang, Y. F. Guo, and W. X. Xu, "Sies: a novel implementation of spiking convolutional neural network inference engine on field-programmable gate array," *Journal of Computer Science and Technology*, vol. 35, no. 2, pp. 475–489, Mar. 2020.

[60] Y. Liu, C. Yang, L. Jiang, S. L. Xie, and Y. Zhang, "Intelligent edge computing for IoT-based energy management in smart cities," *IEEE Network*, vol. 33, no. 2, pp. 111–117, Mar. 2019.

[61] ARM. CMSIS NN software library. [Online]. Available: https://www.keil.com/pack/doc/CMSIS/NN/html/index.html.

[62] A. Sebastian, M. Le Gallo, R. Khaddam-Aljameh, and E. Eleftheriou, "Memory devices and applications for in-memory computing," *Nature Nanotechnology*, vol. 15, no. 7, pp. 529–544, Mar. 2020.

[63] Z. F. Deng, B. B. Wang, Y. L. Xu, T. T. Xu, C. X. Liu, and Z. L. Zhu, "Multi-scale convolutional neural network with time-cognition for multi-step short-term load forecasting," *IEEE Access*, vol. 7, pp. 88058–88071, Jul. 2019.

[64] S. M. J. Jalali, S. Ahmadian, A. Khosravi, M. Shafie-Khah, S. Nahavandi, and J. P. S. Catalão, "A novel evolutionary-based deep convolutional neural network model for intelligent load forecasting," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8243–8253, Dec. 2021.

[65] L. L. Wen, K. L. Zhou, and S. L. Yang, "Load demand forecasting of residential buildings using a deep learning model," *Electric Power Systems Research*, vol. 179, no. 106073, Feb. 2020.

[66] Z. L. Li, Y. Z. Li, Y. Liu, P. Wang, R. Z. Lu, and H. B. Gooi, "Deep learning based densely connected network for load forecasting," *IEEE Transactions on Power Systems*, vol. 36, no. 4, pp. 2829–2840, Jul. 2021.

[67] Y. R. Chen, Z. K. Dong, Y. Wang, J. Su, Z. L. Han, D. Zhou, K. Zhang, Y. S. Zhao, and Y. Bao, "Short-term wind speed predicting framework based on EEMD-GA-LSTM method under large scaled wind history," *Energy Conversion and Management*, vol. 227, no. 113559, Jan. 2021.

[68] C. Yildiz, H. Acikgoz, D. Korkmaz, and U. Budak, "An improved residual-based convolutional neural network for very short-term wind power forecasting," *Energy Conversion and Management*, vol. 228, no. 113731, Jan. 2021.

[69] H. X. Zang, L. L. Cheng, T. Ding, K. W. Cheung, M. M. Wang, Z. N. Wei, and G. Q. Sun, "Application of functional deep belief network for estimating daily global solar radiation: a case study in China," *Energy*, vol. 191, no. 116502, Jan. 2020.

[70] M. Ding, H. Zhou, H. Xie, M. Wu, K. Z. Liu, Y. Nakanishi, and R. Yokoyama, "A time series model based on hybrid-kernel least-squares support vector machine for short-term wind power forecasting," *ISA Transactions*, vol. 108, pp. 58–68, Feb. 2021.

[71] M. W. Ahmad, M. Mourshed, and Y. Rezgui, "Tree-based ensemble methods for predicting PV power generation and their comparison with support vector regression," *Energy*, vol. 164, pp. 465–474, Dec. 2018.

[72] F. Najibi, D. Apostolopoulou, and E. Alonso, "Enhanced performance Gaussian process regression for probabilistic short-term solar output forecast," *International Journal of Electrical Power & Energy Systems*, vol. 130, pp. 106916, Sep. 2021.

[73] A. Ahmadi, M. Nabipour, B. Mohammadi-Ivatloo, A. M. Amani, S. Rho, and M. J. Piran, "Long-term wind power forecasting using tree-based learning algorithms," *IEEE Access*, vol. 8, pp. 151511–151522, Aug. 2020.

[74] S. J. Chai, Z. Xu, Y. W. Jia, and W. K. Wong, "A robust spatiotemporal forecasting framework for photovoltaic generation," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5370–5382, Nov. 2020.

[75] M. M. Islam, Z. Y. Sun, R. W. Qin, W. Q. Hu, H. Y. Xiong, and K. B. Xu, "Flexible energy load identification in intelligent manufacturing for demand response using a neural network integrated particle swarm optimization," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 236, no. 4, no. 1943–1959, Jun. 2022.

[76] K. Tornai, A. Oláh, R. Drenyovszki, L. Kovács, I. Pinté, and J. Levendovszky, "Recurrent neural network based user classification for smart grids," in *Proceedings of 2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference*, 2017, pp. 1–5.

[77] Z. Wen, D. O'Neill, and H. Maei, "Optimal demand response using device-based reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2312–2324, Sep. 2015.

[78] B. Wang, Y. Li, W. Y. Ming, and S. R. Wang, "Deep reinforcement learning method for demand response management of interruptible load," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3146–3155, Jul. 2020.

[79] S. Bahrami, Y. C. Chen, and V. W. S. Wong, "Deep reinforcement learning for demand response in distribution networks," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1496–1506, Mar. 2021.

[80] L. A. Hurtado, E. Mocanu, P. H. Nguyen, M. Gibescu, and R. I. G. Kamphuis, "Enabling cooperative behavior for building demand response based on extended joint action learning," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 127–136, Jan. 2018.

[81] G. Y. Liu, W. J. Yu, and L. Zhu, "Experiment-based supervised learning approach toward condition monitoring of PV array mismatch," *IET Generation, Transmission & Distribution*, vol. 13, no. 7, pp. 1014–1024, Apr. 2019.

[82] I. M. Black, M. Richmond, and A. Kolios, "Condition monitoring systems: a systematic literature review on machine-learning methods improving offshore-wind turbine operational management," *International Journal of Sustainable Energy*, vol. 40, no. 10, pp. 923–946, Mar. 2021.

[83] K. Kudelina, T. Vaimann, B. Asad, A. Rassõlkin, A. Kallaste, and G. Demidova, "Trends and challenges in intelligent condition monitoring of electrical machines using machine learning," *Applied Sciences*, vol. 11, no. 6, pp. 2761, Mar. 2021.

[84] A. Doolgindachbaporn, G. Callender, P. L. Lewin, E. Simonson, and G. Wilson, "Data driven transformer thermal model for condition monitoring," *IEEE Transactions on Power Delivery*, vol. 37, no. 4, pp. 3133–3141, Aug. 2021.

[85] Y. Q. Chen, O. Fink, and G. Sansavini, "Combined fault location and classification for power transmission lines fault diagnosis with integrated feature extraction," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, pp. 561–569, Jan. 2018.

[86] C. Xiao, Z. J. Liu, T. L. Zhang, and X. Zhang, "Deep learning method for fault detection of wind turbine converter," *Applied Sciences*, vol. 11, no. 3, pp. 1280, Jan. 2021.

[87] J. Vives, E. Quiles, and E. García, "AI techniques applied to diagnosis of vibrations failures in wind turbines," *IEEE Latin America Transactions*, vol. 18, no. 8, pp. 1478–1486, Aug. 2020.

[88] Y. P. Liu, S. T. Pei, W. P. Fu, K. Y. Zhang, X. X. Ji, and Z. H. Yin, "The discrimination method as applied to a deteriorated porcelain insulator used in transmission lines on the basis of a convolution neural network," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 24, no. 6, pp. 3559–3566, Dec. 2017.

[89] A. Zollanvari, K. Kunanbayev, S. A. Bitaghsir, and M. Bagheri, "Transformer fault prognosis using deep recurrent neural network over vibration signals," *IEEE Transactions on Instrumentation and Measurement, vol*. 70, pp. 2502011, Sep. 2021.

[90] Y. Du and F. X. Li, "Intelligent multi-microgrid energy management based on deep neural network and model-free reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1066–1076, Mar. 2020.

[91] Q. Z. Zhang, K. Dehghanpour, Z. Y. Wang, F. Qiu, and D. B. Zhao, "Multi-agent safe policy learning for power management of networked microgrids," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1048–1062, Mar. 2021.

[92] Y. W. Shang, W. C. Wu, J. B. Guo, Z. Ma, W. X. Sheng, Z. Lv, and C. R. Fu, "Stochastic dispatch of energy storage in microgrids: an augmented reinforcement learning approach," *Applied Energy*, vol. 261, pp. 114423, Mar. 2020.

[93] T. Y. Chen, S. R. Bu, X. Liu, J. K. Kang, F. R. Yu, and Z. Han, "Peer-to-peer energy trading and energy conversion in interconnected multi-energy microgrids using multi-agent deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 715–727, Jan. 2022.

[94] M. S. Munir, S. F. Abedin, N. H. Tran, Z. Han, E. N. Huh, and C. S. Hong, "Risk-aware energy scheduling for edge computing with microgrid: a multi-agent deep reinforcement learning approach," *IEEE Transactions on Network and Service Management*, vol. 18, no. 3, pp. 3476–3497, Sep. 2021.

[95] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. T. Gu, and V. O. K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018.

[96] A. Gupta, G. Gurrala, and P. S. Sastry, "An online power system stability monitoring system using convolutional neural networks," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 864–872, Mar. 2019.

[97] H. Y. Wang, Q. F. Chen, and B. H. Zhang, "Transient stability assessment combined model framework based on cost-sensitive method," *IET Generation, Transmission & Distribution*, vol. 14, no. 12, pp. 2256–2262, Jun. 2020.

[98] X. Li, Z. G. Yang, P. F. Guo, and J. Z. Cheng, "An intelligent transient stability assessment framework with continual learning ability," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 12, pp. 8131–8141, Dec. 2021.

[99] T. Su, Y. B. Liu, J. B. Zhao, and J. Y. Liu, "Deep belief network enabled surrogate modeling for fast preventive control of power system transient stability," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 315–326, Jan. 2022.

[100] T. J. Wang and Y. Tang, "Transient stability preventive control based on graph convolution neural network and transfer deep reinforcement learning," *CSEE Journal of Power and Energy Systems*, doi: 10.17775 /CSEEJPES.2022.05030.

[101] T. J. Wang and Y. Tang, "Automatic Voltage Control of Differential Power Grids Based on Transfer Learning and Deep Reinforcement Learning", in *CSEE Journal of Power and Energy Systems*, vol. 9, no. 3, pp. 937–948, May 2023.

[102] Q. H. Huang, R. K. Huang, W. T. Hao, J. Tan, R. Fan, and Z. Y. Huang, "Adaptive power system emergency control using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1171–1182, Mar. 2020.

[103] Z. M. Yan and Y. Xu, "A multi-agent deep reinforcement learning method for cooperative load frequency control of a multi-area power system," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4599–4608, Nov. 2020.

[104] S. Y. Wang, J. J. Duan, D. Shi, C. L. Xu, H. F. Li, R. S. Diao, and Z. W. Wang, "A data-driven multi-agent autonomous voltage control framework using deep reinforcement learning," *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4644–4654, Nov. 2020.

[105] Z. H. Wang, H. B. He, Z. Q. Wan, and Y. Sun, "Coordinated topology attacks in smart grid using deep reinforcement learning," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1407–1415, Feb. 2021.

[106] M. Ismail, M. F. Shaaban, M. Naidu, and E. Serpedin, "Deep learning detection of electricity theft cyber-attacks in renewable distributed generation," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3428–3437, Jul. 2020.

[107] Y. Zhang, J. H. Wang, and B. Chen, "Detecting false data injection attacks in smart grids: a semi-supervised deep learning approach," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 623–634, Jan. 2021.

[108] H. Z. Wang, J. Q. Ruan, G. B. Wang, B. Zhou, Y. T. Liu, X. Q. Fu, and J. C. Peng, "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4766–4778, Nov. 2018.

[109] A. J. Abianeh, Y. H. Wan, F. Ferdowsi, N. Mijatovic, and T. Dragičević, "Vulnerability identification and remediation of FDI attacks in islanded DC microgrids using multiagent reinforcement learning," *IEEE Transactions on Power Electronics*, vol. 37, no. 6, pp. 6359–6370, Jun. 2022.

[110] G. M. Gilbert, S. Naiman, H. Kimaro, and B. Bagile, "A critical review of edge and fog computing for smart grid applications," in *Proceedings of the 15th IFIP WG 9.4 International Conference on Social Implications of Computers in Developing Countries*, 2019, pp. 763–775.

[111] W. J. Huo, F. C. Liu, L. Y. Wang, Y. F. Jin, and L. Wang, "Research on distributed power distribution fault detection based on edge computing," *IEEE Access*, vol. 8, pp. 24643–24652, Dec. 2020.

[112] N. Peng, R. Liang, G. H. Wang, P. Sun, C. Y. Chen, and T. Y. Hou, "Edge computing-based fault location in distribution networks by using asynchronous transient amplitudes at limited nodes," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 574–588, Jan. 2021.

[113] S. L. Chen, H. Wen, J. S. Wu, W. X. Lei, W. J. Hou, W. J. Liu, A. D. Xu, and Y. X. Jiang, "Internet of things based smart grids supported by intelligent edge computing," *IEEE Access*, vol. 7, pp. 74089–74102, Jun. 2019.

[114] T. Sirojan, S. B. Lu, B. T. Phung, D. M. Zhang, and E. Ambikairajah, "Sustainable deep learning at grid edge for real-time high impedance fault detection," *IEEE Transactions on Sustainable Computing*, vol. 7, no. 2, pp. 346–357, Apr./Jun. 2022.

[115] J. Tong, H. H. Wu, Y. J. Lin, Y. G. He, and J. Y. Liu, "Fog-computing-based short-circuit diagnosis scheme," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 3359–3371, Jul. 2020.

[116] W. J. Hou, Y. X. Jiang, W. X. Lei, A. D. Xu, H. Wen, and S. L. Chen, "A P2P network based edge computing smart grid model for efficient resources coordination," *Peer-to-Peer Networking and Applications*, vol. 13, no. 3, pp. 1026–1037, Jan. 2020.

[117] D. Gebbran, G. Verbič, A. C. Chapman, and S. Mhanna, "Coordination of prosumer agents via distributed optimal power flow: an edge computing hardware prototype," in *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 2020, pp. 2104–2106.

[118] M. Dabbaghjamanesh, A. Kavousi-Fard, and Z. Y. Dong, "A novel distributed cloud-fog based framework for energy management of networked microgrids," *IEEE Transactions on Power Systems*, vol. 35, no. 4, pp. 2847–2862, Jul. 2020.

[119] S. Y. Wang, X. D. Wang, and W. C. Wu, "Cloud computing and local chip-based dynamic economic dispatch for microgrids," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 3774–3784, Sep. 2020.

[120] T. J. Pu, X. Y. Wang, Y. F. Cao, Z. C. Liu, C. Qiu, J. Qiao, and S. H. Zhang, "Power flow adjustment for smart microgrid based on edge computing and multi-agent deep reinforcement learning," *Journal of Cloud Computing*, vol. 10, no. 1, pp. 48, Sep. 2021.

[121] W. X. Lei, H. Wen, J. S. Wu, and W. J. Hou, "MADDPG-based security situational awareness for smart grid with intelligent edge," *Applied Sciences*, vol. 11, no. 7, pp. 3101, Mar. 2021.

[122] X. Y. Zhang, D. Biagioni, M. M. Cai, P. Graf, and S. Rahman, "An edge-cloud integrated solution for buildings demand response using reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 420–431, Jan. 2021.

[123] X. D. Pang, X. C. Lu, H. Ding, and J. M. Guerrero, "Construction of smart grid load forecast model by edge computing," *Energies*, vol. 15, no. 9, pp. 3028, 2022.

[124] Y. Liao and J. B. He, "Optimal smart grid operation and control enhancement by edge computing," in *Proceedings of 2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2020, pp. 1–6.

[125] I. Galanis, S. S. N. Perala, and I. Anagnostopoulos, "Edge computing and efficient resource management for integration of video devices in smart grid deployments," in *IoT for Smart Grids*, K. Siozios, D. Anagnostos, D. Soudris, and E. Kosmatopoulos, Eds. Cham: Springer, 2019, pp. 115–132.

[126] A. H. Sodhro, S. Pirbhulal, and V. H. C. de Albuquerque, "Artificial intelligence-driven mechanism for edge computing-based industrial applications," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 7, pp. 4235–4243, Jul. 2019.

[127] N. L. Qin, B. Li, D. Li, X. S. Jing, C. Y. Du, and C. Y. Wan, "Resource allocation method based on mobile edge computing in smart grid," *IOP Conference Series: Earth and Environmental Science*, vol. 634, no. 1, pp. 012054, Feb. 2021.

[128] D. Jin, Z. Y. Li, C. Hannon, C. Chen, J. H. Wang, M. Shahidehpour, and C. W. Lee, "Toward a cyber resilient and secure microgrid using software-defined networking," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2494–2504, Sep. 2017.

[129] R. Borgaonkar, I. Anne Tøndel, M. Zenebe Degefa, and M. Gilje Jaatun, "Improving smart grid security through 5G enabled IoT and edge computing," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 18, pp. e6466, Sep. 2021.

[130] S. A. Chaudhry, H. Alhakami, A. Baz, and F. Al-Turjman, "Securing demand response management: a certificate-based access control in smart grid edge computing infrastructure," *IEEE Access*, vol. 8, pp. 101235–101243, May 2020.

[131] Z. H. Wang, D. D. Jiang, F. Wang, Z. H. Lv, and R. Nowak, "A polymorphic heterogeneous security architecture for edge-enabled smart grids," *Sustainable Cities and Society*, vol. 67, pp. 102661, Apr. 2021.

[132] X. N. Zhang, F. Fang, and J. Q. Wang, "Probabilistic solar irradiation forecasting based on variational bayesian inference with secure federated learning," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 11, pp. 7849–7859, Nov. 2021.

[133] F. Sattler, S. Wiedemann, K. R. Müller, and W. Samek, "Robust and communication-efficient federated learning from non-i.i.d. data," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 9, pp. 3400–3413, Sep. 2020.

[134] M. Y. Mehmood, M. Abrar, H. M. Munir, S. F. Hasan,.H. A. U. Muqeet, and N. A. Golilarz, "Edge computing for IoT-enabled smart grid," *Security and Communication Networks*, vol. 2021, pp. 5524025, Jul. 2021.

[135] P. C. M. Arachchige, P. Bertok, I. Khalil, D. X. Liu, S. Camtepe, and M. Atiquzzaman, "Local differential privacy for deep learning," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 5827–5842, Jul. 2020.

[136] X. M. Chang, W. Li, and A. Y. Zomaya, "A lightweight short-term photovoltaic power prediction for edge computing," *IEEE Transactions on Green Communications and Networking*, vol. 4, no. 4, pp. 946–955, Dec. 2020.

[137] L. Lin, X. Guan, Y. Peng, N. Wang, S. Maharjan, and T. Ohtsuki, "Deep reinforcement learning for economic dispatch of virtual power plant in internet of energy," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6288–6301, Jul. 2020.

[138] Y. J. Zhang, X. J. Wang, J. H. He, Y. Xu, F. Zhang, and Y. P. Luo, "A transfer learning-based high impedance fault detection method under a cloud-edge collaboration framework," *IEEE Access*, vol. 8, pp. 165099–165110, Sep. 2020.

[139] Z. W. Niu, M. Z. Reformat, W. H. Tang, and B. N. Zhao, "Electrical equipment identification method with synthetic data using edge-oriented generative adversarial network," *IEEE Access*, vol. 8, pp. 136487–136497, Jul. 2020.

[140] S. H. Zhang, J. Y. Wang, J. Tong, J. Zhang, and M. H. Zhang, "Cloud-edge fusion based abnormal object detection of power transmission lines using incremental learning," *IEEE Access*, vol. 8, pp. 218694–218701, Nov. 2020.

[141] Y. M. Xiang, Z. H. Yi, X. Lu, Z. Yu, D. Shi, C. L. Xu, X. M. Li, and Z. W. Wang, "Distributed frequency emergency control with coordinated edge intelligence," *Electric Power Systems Research*, vol. 199, pp. 107448, Oct. 2021.

[142] A. Nammouchi, P. Aupke, A. Kassler, A. Theocharis, V. Raffa, and M. Di Felice, "Integration of AI, IoT and edge-computing for smart microgrid energy management," in *Proceedings of 2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe*, 2021, pp. 1–6.

[143] L. Zhao, J. M. Li, Q. Li, and F. Y. Li, "A federated learning framework for detecting false data injection attacks in solar farms," *IEEE Transactions on Power Electronics*, vol. 37, no. 3, pp. 2496–2501, Mar. 2022.

[144] H. Y. Wu, P. Chen, W. Li, C. X. Jiang, J. X. Li, and P. Y Zhu, "Fault prediction of power communication networks by federated learning of distributed data," in *Proceedings of the IEEE 6th International Conference on Computer and Communication Systems*, 2021, pp. 605–609.

[145] M. S. Munir, S. F. Abedin, N. H. Tran, and C. S. Hong, "When edge computing meets microgrid: a deep reinforcement learning approach," *IEEE Internet of Things Journal*, vol. 6, no. 5, pp. 7360–7374, Oct. 2019.

[146] W. Li, T. Yang, F. C. Delicato, P. F. Pires, Z. Tari, S. U. Khan, and A. Y. Zomaya, "On enabling sustainable edge computing with renewable energy resources," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 94–101, May 2018.

[147] M. S. Li, N. Cheng, J. Gao, Y. L. Wang, L. Zhao, and X. M. Shen, "Energy-efficient UAV-assisted mobile edge computing: resource allocation and trajectory optimization," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 3, pp. 3424–3438, Mar. 2020.

[148] W. Zhou, L. Xing, J. J. Xia, L. S. Fan, and A. Nallanathan, "Dynamic computation offloading for MIMO mobile edge computing systems with energy harvesting," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 5172–5177, May 2021.

[149] G. Perin, M. Berno, T. Erseghe, and M. Rossi, "Towards sustainable edge computing through renewable energy resources and online, distributed and predictive scheduling," *IEEE Transactions on Network and Service Management*, vol. 19, no. 1, pp. 306–321, Mar. 2022.

Hoay Beng Gooi (Fellow, IEEE) received the B.S. degree in Electrical Engineering from National Taiwan University, Taipei, Taiwan, in 1978, the M.S. degree in Electrical Engineering from the University of New Brunswick, Fredericton, NB, Canada, in 1980, and the Ph.D. degree in Electrical Engineering from The Ohio State University, Columbus, OH, USA, in 1983. He was an Assistant Professor with Lafayette College, Easton, PA, USA, from 1983 to 1985, and a Senior Engineer with Control Data-Energy Management System Division, Plymouth, MN, USA, for about six years before joining Nanyang Technological University (NTU), Singapore, in 1991. He is an Associate Professor with the School of Electrical and Electronic Engineering. Since 2020, he has been the Co-Director of SP Group-NTU Joint Laboratory. His current research interests include microgrid energy management systems dealing with energy storage, condition monitoring, electricity market, and spinning reserve.

Tianjing Wang received a Ph.D. degree from China Electric Power Research Institute, Beijing, China, in 2022. She is currently a research fellow at the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. From Aug. 2021 to Feb. 2022, she was a visiting scholar at the school of Electrical and Electronic Engineering, Nanyang Technological University. Her research interests include deep reinforcement learning applications in power system control, transfer learning and federated learning applications in power systems.

Yong Tang (SM'12) received a Ph.D. degree in Power System and Automation from China Electric Power Research Institute, Beijing, China, in 2002. He is currently a Chair Professor with CEPRI. He is a Fellow of CSEE and a senior member of IEEE. His research interests are power system simulation and analysis, voltage stability and control, load modeling, and simulation.