Review on the Research and Practice of Deep Learning and Reinforcement Learning in Smart Grids

Dongxia Zhang, Xiaoqing Han, and Chunyu Deng

***Abstract*—Smart grids are the developmental trend of power systems and they have attracted much attention all over the world. Due to their complexities, and the uncertainty of the smart grid and high volume of information being collected, artificial intelligence techniques represent some of the enabling technologies for its future development and success. Owing to the decreasing cost of computing power, the profusion of data, and better algorithms, AI has entered into its new developmental stage and AI 2.0 is developing rapidly. Deep learning (DL), reinforcement learning (RL) and their combination–deep reinforcement learning (DRL) are representative methods and relatively mature methods in the family of AI 2.0. This article introduces the concept and status quo of the above three methods, summarizes their potential for application in smart grids, and provides an overview of the research work on their application in smart grids.**

***Index Terms*—Deep learning, deep reinforcement learning, reinforcement learning, smart grid.**

I. INTRODUCTION

ARTIFICIAL intelligence is a comprehensive science and technology, including a huge domain in breadth and width. After decades of development, much progress has been made across the fields of artificial intelligence (AI). In recent years, driven by the increasing amounts of data, there has been a significant emergence in the advancement of AI algorithms and powerful computer hardware, allowing AI to enter into a new evolutionary stage: AI 2.0 [1]. AI 2.0 related technologies are currently in the process of development, and many algorithms are emerging. In their present states, deep learning (DL) and reinforcement learning (RL) are relatively mature; and their combination-deep reinforcement learning (DRL) has won some success with AlphaGo, a game-playing program developed by Google DeepMind. In March 2016, AlphaGo defeated the world Go champion, Sedol Lee, with a 4:1 score, and thereby attracted a new wave of global attention [2].

Manuscript received April 27, 2018; revised July 11, 2018; accepted July 17, 2018; date of publication September 30, 2018; date of current version July 30, 2018.

D. X. Zhang (corresponding author, e-mail: zhangdx@epri.sgcc.com.cn) and C. Y. Deng are with China Electric Power Research Institute, Beijing 100192, China.

X. Q. Han is with the Department of Electrical Engineering, Taiyuan University of Technology, Taiyuan 030024, China.

DOI: 10.17775/CSEEJPES.2018.00520

A. Deep Learning

DL is a subset of machine learning, which originally resulted from a multi-layer Artificial Neural Network (ANN). Strictly speaking, DL has a wider meaning, but in its present state, when talking about DL, we just think of a large deep neural network, that is, deep neural networks. Here, deep typically refers to the number of layers.

There are different structures of DL: Boltzmann Machine (BM), Deep Belief Networks (DBN), Feedforward Deep Networks (FDN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Long-short Term Memory (LSTM) Networks, Generative Adversarial Networks (GAN). Among them, CNN and RNN are the most popular structures. CNN is suitable for dealing with spatial distribution data while RNN has advantages in managing time series data.

DL frameworks are basic underlying architectures, consisting of popular ANN algorithms, providing a stable DL API (Application Program Interface), supporting distributed learning of training models on GPU, TPU, and playing an important role in popularizing DL applications. Mainstream open source DL frameworks include TensorFlow, Caffe/Caffe2, CNTK, MXNet, Paddle-paddle, Torch/PyTorch, Thenano, etc.

Presently, DL has shown its extraordinary ability in many areas, such as image and speech recognition and natural language processing [3], [4].

B. Reinforcement Learning

There are four basic components in RL: agent, environment, reward and action. The objective of RL is for the agent to maximize the reward by taking a series of actions in response to a dynamic environment. RL can support sequential decision making under uncertainty. A typical RL algorithm operates with only limited knowledge of the environment and with limited feedback on the quality of the decisions.

The most popular algorithms of RL include Q-learning, SARSA (State–Action–Reward–State–Action), DQN (Deep Q Net) and DDPG (Deep Deterministic Policy Gradients). More advanced algorithms are NAF (Normalized Advantage Functions), A3C (Asynchronous Advantage Actor-Critic), etc. [5], [6]. RL is often used for robotics, gaming and navigation.

C. Deep Reinforcement Learning

The combination of DL with RL has led to a new field of research, called DRL, which integrates the perception of DL and the decision making of RL. Therefore, DRL can implement a variety of tasks requiring both rich perception of high-dimensional raw inputs and policy control [7]. DRL's success in AlphoGo demonstrates that DRL has become one of the most fascinating research areas in AI.

Many researches on DRL applications are currently being conducted. For example, Google Brain is taking advantage of DRL to train robotic arms to open doors and pick up objects on their own. Uber is trying to use DRL to teach Grand Theft Auto to handle real cars on real roads. As the key algorithms of machine learning, DL and RL have existed in academic circles for many years, but only recently have they entered the industrial world for application oriented research. This is primarily because of the changes or progress in three respects: increasing amounts of data, and the emergence of advanced AI algorithms and powerful computer hardware. Sensors, monitoring systems, smartphones and social media are producing massive data every minute, which brings us into the era of big data. Data deluge accelerates the development of better algorithms. The development of computational power also plays an important role in promoting, especially DL, RL and of course, DRL. Training a DL model without a GPU would be painfully slow in most cases.

Although DRL has proved its value in some practical applications and many experts think DRL has great prospects, some experts have expressed different viewpoints. They think that the present versions of DRL take a long time to converge and learn something meaningful. This restricts the techniques from being used in the real world for real-time learning. This is why this field is having such a great success for example in video games (Atari) or strategy games whose rules are known (Go). They believe that the only way that DRL techniques can be used in a practical way is either pre-training in a simulator or by transferring the learned knowledge between different domains [8].

This article presents an overview of the research work being done on the applications of DL, RL and DRL in a smart grid. Section I briefly introduces the concept of DL, RL and DRL. Section II outlines the application potential of these algorithms in a smart grid. Sections III and IV provide an overview of related literature and introduce some application oriented research work. Finally, Section V contains the conclusion.

II. Application Potential in Smart Grid

Power systems are very complex artificial systems. With the development of the smart grid, high penetration of wind, solar power and customers' active participation have lead smart grids to operate in more uncertain, complex environments. Traditional power system analysis and control decision-making are primarily dependent on physical modeling and numerical calculations. The traditional methods find difficulty in addressing uncertainty and partial observability issues so that they cannot meet the requirements of future development of smart grids. On the other hand, wide spread deployment of Advanced Metering Infrastructures (AMI), Wide Area Monitoring Systems (WAMS), and other monitoring/management systems produce massive data and provide a data basis for algorithm/model training in AI applications. Therefore, DL, RL, and DRL appear to be some of the enabling technologies for the future development and success of smart grids.

First, uncertainties of wind and solar energy bring many challenges for power systems. DL is a potentially powerful tool to improve solar, and wind generation prediction accuracy based on large datasets. Besides, DL, RL and DRL can provide effective solutions for managing flexible sources, including load forecasting, scheduling and dealing, etc.

In an uncertain and complex environment, to ensure secure and stable operations of large scale power systems is one of the biggest challenges that power engineers have to address today. Traditionally, power system operations and decision-making of controls are based on power system numerical calculations based on physical models describing the behavior of power systems. In general, physical models are built according to some assumptions and simplifications and such is the case with power system models. However, the complexity of the power system stability problems along with uncertainties and nonlinearities make the models non-practical or inaccurate. DL and DRL algorithms are believed to significantly improve current control schemes to solve decision and control problems.

Demand Response (DR) is an effective measure to incentivize customers to shift load from peak periods to off-peak periods or to decrease their electricity usage during peak time. With the deployment of smart meters, data reflecting consumers' energy consumption behavior can be collected, and data driven DL and RL algorithms can serve as effective technologies for non-invasive load decomposition, price forecasting, etc., so that consumers can make the right decisions and demand response can be successfully implemented.

In a smart grid, many monitoring systems are installed for equipment supervision. Based on the data from these systems, utilizing DL algorithms, various features can be extracted so that anomalies can be detected, including equipment anomalies, malicious attacks and false data invasions, electricity theft, etc. DL is believed to be of great value in these fields. In smart grids, there are many other technical fields in which DL, RL and DRL have great potential applications. In fact, these algorithms have great prospects in nearly all the technical fields of a smart grid, as shown in the next section.

III. Literature Review

Many papers/dissertations have been published, covering DL, RL and DRL applications in smart grids. Most of them were published since 2016. These application fields cover load/power consumption forecasting, microgrids, demand response, defect/fault detection, cyber security, stability analysis, and nearly all the technical fields of smart grids. In the interest of brevity, this section reviews some of the available literature.

A. Load/demand Forecasting

With the high penetration of solar, and wind power, the scheduling and operation of power systems are faced with the challenges of increasing uncertainty. Therefore, accurate forecasting of energy demands at different levels is important. Although there exist many methods for electricity load forecasting, most of them are based on small datasets without considering the large volumes of data provided by smart meters.

In paper [9], the authors attempt to improve the load forecasting accuracy for ISO New England by trying to use various statistical and AI models. A feedforward neural network is adopted, in which different numbers of hidden layers are chosen. As the number of hidden layers increases, the forecast accuracy improves while the training takes a longer time and the risk of overfitting increases. The research result shows that 100 hidden layers is the best choice.

In many cases, more effective solutions can be obtained by combining a DL algorithm with other methods Some experts are exploring using these types of joint methods for load forecasting. In paper [10] CNN and K-Means algorithms are used jointly to forecast hourly electricity loads. A K-means algorithm is used to cluster large datasets, which contain more than 1.4 million of load records, into subsets and the obtained subsets are used to train the CNN. Experimental results demonstrate that the proposed method is effective. Paper [11] proposes one DBN embedded with parametric Copula models to forecast the hourly load, based on one year data in an urban area in Texas, United States. Experimental results show the proposed method can give more accurate predictions than classical ANN, SVR (Support Vector Regression), and ELM (Extreme Learning Machine). In paper [12], DNN (Deep Neural Network) is applied for day-ahead load prediction. A 90-day Iberian market dataset is used to train multiple combinations of activation functions of both single and double-layer neural networks. The results indicate that the combination of an Exponential linear unit (ELU) with ELU has better performances than other combinations when evaluated against MAPE (Mean Absolute Percentage Error) values for week day datasets; while for weekend data sets, the ReLU-ReLU (Rectifier linear unit) combination outperforms the other combinations.

Demand forecasting at the individual building level is a new and important topic since it can help carry out demand response locally. Meanwhile, smart meters collect energy consumption data at building and individual site levels, which enable DL based load/demand forecasting. Papers [13] and [14] try using Conditional Restricted Boltzmann Machines (CRBM) and Factored Conditional Restricted Boltzmann Machines (FCBRM) to forecast building level loads. Paper [15] proposes two LSTM based neural network architectures to predict one hour and one minute time-step loads. It shows that a standard LSTM architecture cannot accurately forecast one minute load while the S2S LSTM-based architecture achieves excellent performance in both cases. In another paper [16], the same authors try to use CNN for load forecasting at an individual building level based on historical loads. Results obtained from the CNN are compared with those obtained by S2S based LSTM, FCRBM, "shallow" ANN and SVM (Support Vector Machine) for the same dataset and it shows that CNN outperforms SVR while producing comparable results to the ANN and DL methodologies.

It is important to incorporate the weather data into learning algorithms since weather has a direct effect on the energy consumption. Paper [17] employs DL to predict the energy consumption and power generation together with the weather forecasting numerical simulation. The proposed methods are validated on a small scale decentralized system built inside the University campus.

Paper [18] applies DNN and other machine learning techniques to predict short-term load based on smart meter data. Papers [19] and [20] use DBN for load forecasting.

From these references, it can be concluded that: all kinds of DL algorithms are viable candidates for giving more accurate load forecasting than traditional methods. DL algorithms can deal with large volumes of data from smart meters and other data sources. However, their performances need to be further tested based on different datasets. Maybe for different databases, the answer is different. How to choose the number of layers and parameters are still open questions.

B. Microgrid

Microgrid has renewable energy, EVs and all types of customers included in it, so its planning and operations are faced with challenges from uncertainty because of the difficulty in accurate prediction of future electricity consumption and renewable power generation. DL, RL and DRL are helpful for deriving the solutions of decision-making for microgrids. DRL can deal with the problem of partial observability. Paper [21] applies DRL in the operation of microgrids in the environment of partial observation since a limited amount of data is available. To address the partial observation problem, a specific DL architecture has been designed to extract knowledge from past consumption and production time series as well as any available forecasts. The proposed approach proves to be effective based on the case study of a residential customer located in Belgium.

Energy management is one of the core problems for microgrids. Paper [22] proposes an online dynamic energy management system (DEMS) for a microgrid by combining evolutionary adaptive dynamic programming with an RL framework, in order to realize optimal or near-optimal DEMS in both grid-connected and islanded operation modes.

The planning and operation of energy storage in microgrids is a difficult issue. Papers [23] and [24] apply RL in controlling an energy storage device in a microgrid to maximize self-consumption of the energy production from the local PV system and to minimize the electricity cost and dependency on the local utility grid. In some cases, a variety of energy storage systems (ESSs) with different operational characteristics are used simultaneously in a microgrid to improve its flexibility. In paper [25], an RL strategy is proposed to control the charging and discharging periods of the different battery systems to improve system efficiency.

It is expected that future power grid will have multiple microgrids which need to be coordinated with each other; therefore, it is crucial to introduce trading into microgrids. RL can provide effective solutions for this issue especially in the case of incomplete information. Paper [26] presents a new energy trading framework using RL for each microgrid to choose a strategy to trade the energy in an independent market so as to maximize its average revenue.

Usually, for the microgrid, the service provider is responsible for purchasing electric energy from the utility company and selling it to the customers. However, both the service provider and the customers are faced with challenges due to incomplete information acquisition and the various types of uncertainties in the microgrid. To address this issue, paper [27] develops RL algorithms to enable each service provider and each customer to learn its strategy without a priori information about the microgrid. The proposed RL-based dynamic pricing algorithm proves to be effective without a priori information about the system dynamics and the proposed energy consumption scheduling algorithm can help customers to make reasonable decisions.

Micro grids are developing rapidly all over the world. Due to the co-existing of multi-energy sources, and multiple stakeholders in micro grid, it is necessary to address the challenges from incomplete observability and uncertainty. Reinforcement learning has become an effective tool in energy management, energy dealing, ESSs management, and operational control in micro grids. In the future, deep learning and reinforcement learning can form the technical solutions for multiple micro grid coordination.

C. Demand Response

Identifying and predicting energy flexibility on the demand side is crucial for implementing demand response. At the same time, the latest smart meter applications allow us to conveniently monitor customer side power consumption levels in real time and analyze customers' consumption behavior by means of Nonintrusive Load Monitoring (NILM). By using NILM, it can be determined which appliances are being used and their individual consumption.

Dissertation [28] applies DL to identify the flexibility of loads, and to provide references for demand response. Paper [29] uses Factored Four Way Conditional Restricted Boltzmann Machines to identify and predict flexibility in real time. Paper [30] uses RNN to classify consumers and proves that the proposed algorithms can have better performances than the existing methods; in some cases the new calculated rates can reach nearly 100%.

Actually, direct load control and demand response are high dimension control problems faced with the challenges of partial observability and randomness. In papers [31]–[33], RL is applied for the setting of residential demand response at the device level while in papers [32] and [34], RL is used in an aggregate-and-dispatch setting with thermostatically controlled loads.

Paper [35] uses deep Q-learning and Deep policy gradient to make decisions at both the building level and the aggregated level. It shows that deep policy gradient outperforms deep Q-learning in on-line scheduling of energy resources at both levels, although both methods are able to successfully perform either the minimization of the energy cost or the flattening of the net energy profile. Paper [36] applies Q-learning-based approximate dynamic programming into demand response schemes to support a customer's real-time decision-making. Paper [37] employs RL to cluster the customers based on their capability for implementing curtailments once they receive demand response signals. The RL approach helps the retailers in their decision making. Considering the fact that some demand response applications, e.g., a heat-pump thermostat, are influenced by weather and climate, paper [38] incorporates weather forecast data into a well-known batch RL technique, and fitted Q-iteration to support demand response. The research results indicate that the proposed extension of fitted Q-iteration can improve the performance of standard fitted Q-iteration by 27%. Similarly, paper [39] uses batch RL, and fitted Q-iteration to provide a control policy for a thermostat controller; as a result, the total cost of energy consumption of the electric water heater can be reduced by 15%.

Demand response has close connections with customer side energy management. Paper [40] uses batch RL to realize home energy management which can autonomously define a policy to support decision making for selling or storing surplus energy in smart homes based on historical data for energy prices, energy generation, consumer demand and characteristics of storage systems. Paper [41] also uses Q-learning algorithms in an energy management system for residential demand response, the impact of future energy prices and possible future consumer device selections are also considered.

Some scholars make comparisons between RL and other methods in the field of demand response. Paper [42] proposes a simple and unique tree-like MDP (Markov Decision Process) structure for appliance scheduling, considering all possible search paths of all appliances to be activated. Compared with Q-learning in a smart-home scenario, the proposed approach can enable customers to make faster decisions.

Also, it is important and difficult to design fair compensation mechanisms for demand response (DR). Paper [43] uses an RL algorithm in a demand response program for payments.

How to solve the problem with partial observability in demand response is an open question. Paper [44] uses a neural network with convolutional layers, within a fitted Q-iteration, in a realistic demand response program with partial observability.

Dynamic pricing is an effective tool for encouraging customers to participate in demand response and energy trading. However, in practice, due to the lack of information about the customers' time varying load demand and energy consumption patterns and the volatility of electricity prices in the wholesale market, the implementation of dynamic pricing is highly challenging.

Paper [45] presents an RL-based dynamic pricing algorithm for efficient dynamic pricing without requiring the perfect information about the system dynamics a priori. Paper [46] uses RL algorithms to seek the Nash equilibrium (NE) for the constrained energy trading game among players with incomplete information. Paper [47] explores using Q-learning algorithms to develop pricing strategies for Broker Agents in Smart Grid markets. Paper [48] uses RL for demand response incorporating an automatic generation control scheme.

Demand response is a complicated topic involving appropriate incentive mechanisms, customer's behavior in energy consumption and their response to incentive pricing in relation with their social and psychological attributes, and other outside effects including climate, weather, and policies. It is hard or even impossible to constitute its physical model, so it is suitable to use a data driven method to manage it.

DL, as a data driven approach, can be used to extract the load's characteristics, and RL, also a model free method, can support aggregator, customers and grid operators to make decisions for demand response.

The above literature made various efforts in these areas, but how to analyze flexibility from the customer side, control load directly, and make decisions for all the stakeholders in demand response, especially under the partial observability environments, are still open questions.

D. Defect/Fault Detection of Electrical Equipment

Adequate fault/defect detection of electrical equipment is vitally important in order to ensure reliable power system operations. In a typical power system, many sensors and monitoring systems are installed and gradual changes are analyzed. Because of the complexity of recorded data, however, defects or faults at an early stage cannot be easily recognized. In the following literature, DL is employed to monitor the states of three important components in power networks: insulators, transformers and transmission lines.

Paper [49] proposes taking advantage of high-level discriminative CNNs to extract the features of the insulators and identify their defects. The experimental results show that the proposed method can achieve an accuracy of 93%.

Considering that oil chromatography online-monitoring data is unlabeled during power transformer failure and therefore traditional diagnosis methods often fail to make full use of those unlabeled fault samples in judging transformer fault types, paper [47] establishes a corresponding classification method based on DL. On this basis, a new fault diagnosis method for power transformers is further proposed, in which a large number of unlabeled data from oil chromatogram online monitoring devices and a small number of labeled data from dissolved gas-in-oil analysis (DGA) are fully used in the training process. Testing results from engineering examples indicate that the proposed method is better than that of three other proposed methods: radio, BP neural network and SVM.

Paper [50] proposes a fault diagnosis method of power transformer based on deep neural network. A large number of unlabeled data from oil chromatogram on-line monitoring devices and a small number of labeled data from dissolved gas-in-oil analysis are fully used in training process. Testing results indicate that the diagnosis performance is better than that of three radio, BP neural network and SVM. In order to improve the recognition of power line fault detection, paper [51] proposes one modeling method based on the sparse self-encoding neural network, in which normalized sub-band energy of wavelet decomposition is used as the characteristic parameters for the DL neural network. Then the characteristics of the fault signal are trained to construct the DL structure. The simulation experiment based on IEEE 34 shows that the fault recognition rate exceeds 99%.

Deep learning is one effective means for equipment defect detection and the above references prove this. However, some problems need to be resolved such as how to solve the small sample learning problems, how to identify the small differences between normal conditions and pre-faulted conditions and how to satisfy the need for real time defects identification.

E. Power System Analysis and Control

Disturbance identification, stability assessment, and emergency control are fundamental to ensure the reliability and security of the power system. Wide Area Measurement Systems (WAMS) provide massive volumes of data. Therefore, how to get knowledge from this data, preferably automatically, is an actual challenge for system operators.

Dissertation [52] proposes applying DL frameworks, including Multilayer Perceptions, DBN and CNN, to perform automatic disturbance classification using a set of measurements from several PMUs-installed in the low voltage grid of an interconnected system. The results demonstrate that CNN outperforms the others in terms of classification accuracy.

Paper [53] proposes a DBN based transient stability assessment method, in which DBN is utilized to map the original feature space to a representation space, where the stable cases can be linearly separated from the unstable cases.

Paper [54] uses a multilevel DL model to support power system emergency management which uses the modified hidden Markov model to generate the control input for global optimization.

In paper [55], DRL is used for decision-making for emergency control strategies. Specifically, the CNN is used to extract features for the power system during the transient process and the double Q model and dueling Q model are used to improve the performance from Q-Learning and calculate the Q value, and the control strategy can be obtained by comparing the Q value.

In recent years, power systems are becoming larger and more complex and are being penetrated by more and more new elements, such as wind, solar power, flexible loads and electric vehicles. All these elements result in strong coupling and high uncertainty in the power system. Under such conditions, model based methods, which establish mechanism models with assumptions and simplifications as essential preconditions, are questionable. DL and RL, DRL, as model-free methods, are considered valuable alternatives or additions to model-based methods. On the other hand, since DL and RL have the characteristics of non-interpretability, whether such black box methods are suitable for power system analysis and control is questionable. Some experts suggest that combining model based methods and data driven methods will improve the performance of analysis and control. There is a long way to go in this area.

F. Cyber Security

The integration of ICT into power system infrastructure can effectively improve the quality of the monitoring and control smart grids, on the other hand, the dependence on ICT also increases power systems' vulnerability to malicious attacks. False data injection (FDI) is becoming a severe threat to power systems. How to solve the issue has become a hot topic in recent years.

Paper [56] uses Q-learning to monitor topology changes and analyzes the vulnerability of electrical power grids in sequential topological attacks. The vulnerable sequences leading to critical blackouts in the system can then be found. Paper [57] proposes DL-based algorithms to detect FDI and power theft in real-time. High-dimensional temporal behaviors of the unobservable FDI attacks are featured by using CDBN to detect the potential FDI attacks based on real-time measurements. Paper [58] uses DL to probe and detect the data corruption by analyzing the real-time measurement data from the geographically distributed Phasor Measurement Units (PMUs).

Cyber security of smart grids is one of the most important application areas of DL and RL. Some institutions have become initiated in these kinds of research, see IV in this paper for further reference.

G. Renewable Energy Generation Prediction

The prediction of renewable energy generation output is important to improve their integration in the power grid by dealing with their uncertain and intermittent characteristics.

Paper [59] uses multilayer perceptron neural networks in combination with a multi-objective genetic algorithm, and ELM combined with the nearest neighbors' approach, for short-term wind speed prediction as observed for the region of Regina in Saskatchewan, Canada. Both approaches have good prediction precision.

Recently, there has been a great deal of research on predicting wind ramps. However, the present wind power forecasting methodologies have disadvantages in different weather conditions. Paper [60] uses a Multi-Layer Perceptron Neural Network (MLP-NN) for wind forecasts, using a Weather Research and Forecasting Model (WRF) model as input.

Paper [61] proposes a Long Short-Term Memory (LSTM) based wind power prediction model, which is advanced and practical in the field of wind power prediction.

These papers show that DL methods have been increasingly applied to renewable generation prediction. Since DL has the characteristics of flexibility, self-adaptive learning abilities, and the relaxation of the need of physical and phenomenological assumptions, it is expected that prediction accuracy can be enhanced by combining DL methods with more data sources.

H. Others

There are some other papers which apply DL or RL into other technical branches of power systems.

Thesis [62] uses RNNs to coordinate the response of DER with the power grid. Paper [63] employs DL for a power quality (PQ) event classification, taking advantage of the use of DL in image-file-classification. Paper [64] proposes a neural network based reinforcement learning method for the optimal commitment in the presence of PV sources. Paper [65] proposes a distributed RL collaborative consensus algorithm in AGC in interconnected power grids.

Management of the charging/discharging of EVs faces big challenges since the information on charging flexibility of EVs cannot be known beforehand, including numerous details about each EV (e.g., plug-in times, power limitations, battery size, power curve, etc.). To cope with this challenge, paper [66] uses batch RL to identify the EV charging behavior and then define a cost-effective day ahead consumption plan based on the learned behavior.

Paper [67] introduces a novel Q-learning mechanism to support the feeder automation in decision making for restoration. Paper [68] focuses on implementing a dimensional Q-learning (DQL) for reactive power optimization with discrete control variables. The experimental studies show that DQL is able to optimize the reactive power dispatch and safety margin with greater advantages than the other two popular algorithms.

IV. Application Oriented Research

There are two examples of the next generation of AI technologies [69] which are application oriented. The first one is the patented technology GridSense developed by Alpiq. GridSense is equipped with several self-learning intelligent algorithms, which continually measure parameters such as grid load, power consumption and generation and incorporate weather forecasts and electricity tariffs into their calculations. The algorithms also determine how the different power consumers behave. GridSense has been installed for field tests in the grid area at Biel-Benken, EBM, in Switzerland since 2015. The aim of this field test is to see how well GridSense performs in a modern single-family home. The technology avoids peak loads in the power grid and stabilizes the distribution grid through the efficient use of power consumers. As decentralized production increases, the distribution grid will be put under ever increasing pressure. By optimizing power flows, the amount of new grid construction and expansion can be minimized and costs can be reduced.

The second application-oriented project is DARPA's Rapid Attack Detection, Isolation and Characterization Systems (RADICS) [70]. By means of RADICS, cyber-attacks to power-grid infrastructures can be detected as early as possible by distinguishing between routine outages and actual attacks using AI technologies. RADICS aims to develop advanced anomaly-detection systems with high sensitivity and low false positive rates, based on the analyses of the power grid's dynamics.

V. Conclusion

DL, RL and DRL have attracted a great deal of attention all over the world in recent years. Unlike traditional AI algorithms, such as expert system, fuzzy logic, and neural networks, which are academic driven, DL, RL and DRL are application oriented, and are being promoted by high tech companies. Their combination with available, massive datasets, and inexpensive parallel computing has made them possible to find actual applications.

Smart grids are one of the greatest potential application areas of DL, RL and DRL. Great efforts have been made on researches on their application in smart grids. A lot of papers have been published in this area and most of them were published in the most recent three years. This article only introduces some of them.

From the papers the article refers to, DL, RL and DRL can be used in many technical fields of smart grids. From the technological viewpoint, they can be applied in prediction, anomaly detection, decision-making support for control, etc. For business, they can be used in renewable generation prediction, defect/fault detection of equipment, security assessment and control, cyber security defense, demand response, load forecasting, nearly covering all the technical fields of smart grids. The distribution of the papers are illustrated in Table I.

Indeed, these algorithms have great application potential in the future for smart grids, but at present the related researches are still at their initial stages. Many problems need to be studied in depth, for example, which structure of deep learning is preferable for a specific use case? How many layers are suitable? How to solve the problem of partial observability, and small samples? In addition, DL has the shortcoming of non-interpretable and this can impact the creditability of results, leading to security concerns for the power system operator.

TABLE I
THE DISTRIBUTION OF THE PAPERS

Fields	No. of Papers	Technical Solutions (Challenges to face)
Load Prediction	[9]–[20]	To improve prediction accuracy based on big datasets from AMI and weather forecast systems, etc.
Microgrid	[21]–[27]	To deal with energy management, operational control, and energy dealing under uncertainty and partial observation.
Demand Response	[28]–[48]	Flexibility analysis from the customer's side, NILM, direct load control.
Anomaly Detection	[47], [49] [50]	Anomaly detection of electrical equipment based on data from monitoring systems including unmanned aerial vehicles, inspection robots, etc.
Power System Analysis and Control	[52], [55]	Data driven analysis and control independent of the physical model based on data from WAMS and other data sources to address the uncertainty and complexity of power systems.
Cyber Security	[56]–[58]	Detection of malicious attacks and false data injection.
Renewable Generation Prediction	[59], [60]	To improve prediction accuracy based on big datasets from multiple sources.

REFERENCES

[1] Y. H. Pan, "Heading toward Artificial Intelligence 2.0," *Engineering*, vol. 2, no. 4, pp. 409–413, Dec. 2016.

[2] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre et al., "Mastering the game of Go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.

[3] Z. J. Sun, L. Xue, Y. M. Xu, and Z. Wang, "Overview of deep learning," *Application Research of Computers*, vol. 29, no. 8, pp. 2806–2810, Aug. 2012. (in Chinese).

[4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

[5] R. S. Sutton, and A. G. Barto, *Reinforcement learning – an introduction*. Cambridge, MA: MIT Press, 1998.

[6] X. S. Chen, and Y. M. Yang, "Reinforcement learning: survey of recent work," *Application Research of Computers*, vol. 27, no. 8, pp. 2834–2838, Aug. 2010. (in Chinese).

[7] Q. Liu, J. W. Zhai, Z. Z. Zhang et al., "A survey on deep reinforcement learning," *Chinese Journal of Computers*, vol. 40, no.1, pp. 1–28, Jan. 2017. (in Chinese).

[8] A. Irpan. (Apr. 5, 2018). What's hot in AI: Deep reinforcement learning? [Online]. Available: https://venturebeat.com/2018/04/05/whats-hot-in-ai-deep-reinforcement-learning

[9] L. Voynichka. Machine learning for the smart grid. [Online]. Available: http://cs229.stanford.edu/proj2014/Iliana%20Voynichka,%20Machine%20Learning%20for%20the%20Smart%20Grid.pdf.

[10] X. S. Dong, L. J. Qian, and L. Huang, "Short-term load forecasting in smart grid: a combined CNN and K-means clustering approach," in *IEEE International Conference on Big Data and Smart Computing (BigComp)*, Jeju, Korea, Feb. 13–16, 2017.

[11] Y. S. He, J. H. Deng, and H. J. Li, "Short-term power load forecasting with deep belief network and copula models," in *20179th International Conference on Intelligent Human-Machine Systems and Cybernetics*, Hangzhou, China, Aug. 26–27, 2017.

[12] T. Hossen, S. J. Plathotam, R. K. Angamuthu et al., "Short-term load forecasting using deep neural networks (DNN)," presented at *2017 North American Power Symposium (NAPS)*. [Online]. Available: https://engineering.und.edu/electrical/faculty/prakash-ranganathan/p232.pdf.

[13] V. Mnih, H. Larochelle, and G. Hinton, "Conditional restricted Boltzmann machines for structured output prediction," in *Proceedings of the International Conference on Uncertainty in Artificial Intelligence*, Barcelona, Spain, Jul. 14–17, 2011.

[14] G. W. Taylor, G. E. Hinton, and S. T. Roweis, "Two distributed-state models for generating high-dimensional time series," *Journal of Machine Learning Research*, vol. 12, pp. 1025–1068, Mar. 2011.

[15] D. L. Marino, K. Amarasinghe and M. Manic, "Building energy load forecasting using deep neural networks," in *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, Yokohama, Japan, Oct. 23–26, 2016.

[16] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," in *The 26th IEEE International Symposium on Industrial Electronics (ISIE)*, Edinburgh, Scotland, UK, Jun. 19–21, 2017.

[17] T. Sogabe, H. Ichikawa et al., "Optimization of decentralized renewable energy system by weather forecasting and deep machine learning techniques," in *The 2016 IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia)*, Melbourne, Australia, Nov. 28–Dec.1, 2016.

[18] S. Hosein and P. Hosein, "Load forecasting using deep neural networks," presented at *2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. [Online]. Available: http://www.rndrepository.com/smartgriduwi/files/smartgridcomm_2016.pdf.

[19] A. Dedinec, S. Filiposka, A. Dedinec, and L. Kocarev, "Deep belief network based electricity load forecasting: An analysis of the Macedonian case," *Energy*, vol. 115, part 3, pp. 1688–1700, Nov. 2016.

[20] X. Qiu, L. Zhang, Y. Ren, P. N. Suganthan, and G. Amaratunga, "Ensemble deep learning for regression and time series forecasting," in *2014 IEEE Symposium on Computational Intelligence in Ensemble Learning (CIEL)*, 2014, pp. 1–6.

[21] V. Françoislavet, D. Taralla, D. Ernst et al., "Deep reinforcement learning solutions for energy microgrids management," presented at *European Workshop on Reinforcement Learning (EWRL 2016)*. [Online]. Available: https://orbi.uliege.be/handle/2268/203831

[22] G. K. Venayagamoorthy, R. K. Sharma, P. K. Gautam et al., "Dynamic energy management system for a smart microgrid," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no.8, pp. 1643–1656, Aug. 2016.

[23] B. V. Mbuwir, F. Ruelens, F. Spiessens et al., "Battery energy management in a microgrid using batch reinforcement learning," in *International workshop of Energy-Open 2017*, Enschede, the Netherlands, May 18–19, 2017.

[24] B. V. Mbuwir, F. Ruelens, F. Spiessens et al. (2017, Nov.). Reinforcement learning-based battery energy management in a solar microgrid. *Energies.* [Online]. Available: https://energy-open.nl/local/abstracts/012_mbuwir.pdf.

[25] X. Qiu, T. A. Nguyen and M. L. Crow, "Heterogeneous energy storage optimization for microgrids," *IEEE Transactions on Smart Grid*, vol. 7, no.3, pp.1453–1461, May 2016.

[26] H. W. Wang , T. W. Huang, X. F. Liao et al., "Reinforcement learning in energy trading game among smart microgrid," *IEEE Transactions on Industrial Electronics*, vol. 63, no.8, pp. 5109–5119, Aug. 2018.

[27] B. G. Kim, Y. Zhang, M. V. D. Schaar et al., "Dynamic pricing and energy consumption scheduling with reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2187–2198, Sep. 2016.

[28] Pedro Paulo Marques do Nascimento, "Application of deep learning techniques on HILM," M.S. thesis, COPPE/UFRJ, Brazil, 2016.

[29] D. C. Mocanu, E. Mocanu, P. H. Nguyen et al., "Big IoT data mining for real-time energy disaggregation in buildings," in *IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016*, Budapest, Hungary, Oct. 9–12, 2016.

[30] K. Tornai, A. Olah, R. Drenyovszki et al., "Recurrent neural network based user classification for smart grids," in *IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, Arlington, VA ,Washington D.C., 2017.

[31] 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.

[32] F. Ruelens, B. J. Claessens, S. Vandael, S. Iacovella, P. Vingerhoets, and R. Belmans, "Demand response of a heterogeneous cluster of electric water heaters using batch reinforcement learning," in *2014 Power Systems Computation Conference*, Wroclaw, Poland, Aug. 18–22, 2014.

[33] G. Costanzo, S. Iacovella, F. Ruelens, T. Leurs, and B. Claessens, "Experimental analysis of data-driven control for a building heating system," *Sustainable Energy Grids and Networks*, vol. 6, pp. 81–90, 2016.

[34] E. C. Kara, M. Berges, B. Krogh, and S. Kar, "Using smart devices for system-level management and control in the smart grid: A reinforcement learning framework," in *2012 IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, Taiwan, China, Nov. 5–8, 2012.

[35] E. Mocanu, D. C. Mocanu, P. H. Nguyen et al., "On-line building energy optimization using deep reinforcement learning," *IEEE Transactions on Smart Grid,* to be published.

[36] D. Li, and S. K. Jayaweera, "Machine-learning aided optimal customer decision for an interactive smart grid," *IEEE Systems Journal*, vol. 9, no. 4, pp.1529–1540, Dec. 2015.

[37] S. Ahmed and F. Bouffard, "Building load management clusters using reinforcement learning," in *2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON),* Vancouver, BC, Canada, Oct. 3–5, 2017.

[38] B. J. Claessens, P. Vrancx, and F. Ruelens, "Convolutional neural networks for automatic state-time feature extraction in reinforcement learning applied to residential load control," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3259–3269, Jul. 2018.

[39] F. Ruelens, B. J. Claessens, S. Quaiyum et al., "Reinforcement learning applied to an electric water heater: from theory to practice," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3792–3800, Jul. 2018.

[40] H. Berlink and A. H. R. Costa, "Batch reinforcement learning for smart home energy management," in *Proceedings of the Twenty-Four International Joint Conference on Artificial Intelligence (IJCAI 2015),* Buenos Aires, Argentina, Jul. 25–31, 2015.

[41] D. O' Neill, M. Levorato, A. Goldsmith et al., "Residential demand response using reinforcement learning," in *2010 First IEEE Conference on Smart Grid Communications (SmartGridComm)*, Gaithersburg, MD, USA, Oct. 4–6, 2010.

[42] N. Chauhan, N. Choudhary and K. George, "A comparison of reinforcement learning based approaches to appliances scheduling," in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, India, Dec. 14–17, 2016.

[43] G. O' Brien, A. E. Gamal, and R. Rajagopal, "Shapley value estimation for compensation of participants in demand response programs," *IEEE Transactions on Smart Grid*, vol. 6, no. 6, pp. 2837–2844, Nov. 2015.

[44] F. Ruelens, B. J. Claessens, S. Vandeal et al., "Residential demand response of thermostatically controlled loads using batch reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 8, no.5, pp. 2149–2159, Sep. 2017.

[45] B. G. Kim, Y. Zhang, M. V. D. Schaar, and J. W. Lee, "Dynamic pricing for smart grid with reinforcement learning," *2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS),* Toronto, Ontario, Canada, Apr. 27–May 2, 2014.

[46] H. W. Wang, T. W. Huang, X. F. Liao et al., "Reinforcement learning for constrained energy games with incomplete information," *IEEE Transactions on Cybernetics*, vol. 47, no. 10, pp. 3404–3416, Oct. 2017.

[47] P. P. Reddy and M. M. Veloso, "Strategy learning for autonomous agents in smart grid markets," *Proceedings of the Twenty-Second international joint conference on Artificial Intelligence*, vol. 2, pp. 1446–1451, Jul. 2011.

[48] D. Jay and K. S. Swarup, "Demand response based automatic generation control in smart grid deregulation market," in *IEEE 6th International Conference onPower Systems (ICPS)*, New Delhi, India, Mar. 4–6, 2016.

[49] Z. B. Zhao, X. Q. Fan, G. Z. Xu et al., "Multi-patch deep features for power line insulator status classification from aerial images," *IEEE International Joint Conference on Neural Networks (IJCNN)*, New Delhi, India, Jul. 24–29, 2016.

[50] X. Shi and Y. L. Zhu, "The application of deep neural network in power transformer fault diagnosis," *Electric Power Construction,* vol. 36, no.12, pp.116–122, Dec. 2015. (in Chinese).

[51] X. Peng, F. L. Pan, Y. C. Liang et al., "Fault detection algorithm for power distribution network based on sparse self-encoding neural network," in *2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)*, Changsha, China, May 27–28, 2017.

[52] P. E. A. Cardoso, "Deep learning applied to PMU data in power systems," Ph.D. thesis, Faculdade De Engenharia Da Universidade Do Porto, 2017.

[53] W. Hu, L. Zheng, Y. Min et al., "Research on power system transient stability assessment based on deep learning of big data technique," *Power System Technology*, vol. 41, no. 10, pp. 3140–3146, Oct. 2017. (in Chinese).

[54] X. Z. Wang, J. Zhou, Z. L. Huang et al., "A multilevel deep learning method for big data analysis and emergency management of power system," *IEEE International Conference on Big Data Analysis (ICBDA),* Hangzhou, China, Mar. 12–14, 2016..

[55] W. Liu, D. X. Zhang, X. Y. Wang et al., "A decision making strategy for generating unit tripping under emergency circumstances based on deep reinforcement learning," *Proceedings of the CSEE*, vol. 38, no. 1, pp.109–119, Jan. 2018. (in Chinese).

[56] J. Yan, H. B. He, X. N. Zhong and Y. F. Tang, "Q-learning-based vulnerability analysis of smart grid against sequential topology attacks," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp.200–210, Jan. 2017.

[57] Y. B. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: a deep learning-based intelligent mechanism," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2505–2516, Sep. 2017.

[58] J. Wei and G. J. Mendis, "A deep learning-based cyber-physical strategy to mitigate false data injection attack in smart grids," *Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG),* Vienna, Austria, Apr. 12, 2016.

[59] B. Ak, O. Fink and E. Zio, "Two machine learning approaches for short-term wind speed time-series prediction," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no.8, pp.1734–1747, Aug. 2016.

[60] Y. Q. Li, C. B. Dai, T. X. Wang, Z. Zhou et al., "Separate wind power and ramp predictions based on meteorological variables and clustering method," in *2016 IEEE 6th International Conference on Power Systems (ICPS)*, New Delhi, India, Mar. 4–6, 2016.

[61] X. Y. Qu, X. N. Kang, C. Zhang et al., "Short-term prediction of wind power based on deep long short-term memory," in *2016 IEEE PES Asia-Pacific Power and Energy Conference*, Xi'an, China, Oct. 25–28, 2016.

[62] E. M. Avila, "Artificial recurrent neural networks for the distributed control of electrical grids with photovoltaic electricity," Ph.D. thesis, ETSIT (UPM), Spain, 2016.

[63] E. Balouji and O. Salor, "Classification of power quality events using deep learning on event images," in *2017 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA 2017)*, Shahrekord, Iran, Apr. 19–20, 2017.

[64] E. A. Jasmin, T. P. I. Ahamed, T. Ramani, "A function approximation approach to reinforcement learning for solving unit commitment problem with photo voltaic sources," in *2016 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, Trivandrum, India, Dec. 14–17, 2016.

[65] Q. Li, X. S. Zhang, Z. N. Pan et al., "Decentralized reinforcement learning collaborative consensus algorithm for generation dispatch in virtual generation tribes," in *2015 IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC),* Chongqing, China, Dec. 19–20, 2015.

[66] S. Vandael, B. Claessens, D. Ernst et al., "Reinforcement learning of heuristic EV fleet charging in a day-ahead electricity market," *IEEE Transactions on Smart Grid*, vol. 6, no. 4, pp. 1795–1805, Jul. 2015.

[67] M. J. Ghorbani, M. A. Choudhry, and A. Feliachi, "A multiagent design for power distribution systems automation," *IEEE Transactions on Smart Grid,* vol. 7, no. 1, pp. 329–339, Jan. 2016.

[68] X. Y. Shang, M. S. Li, T. Y. Ji et al., "Discrete reactive power optimization considering safety margin by dimensional Q-Learning," in *2015 IEEE Innovative Smart Grid Technologies Asia (ISGT ASIA),* Bangkok, Thailand, Nov. 3–6, 2015.

[69] https://www.darpa.mil/program/rapid-attack-detection-isolation-and-characterization-systems

[70] http://www.lidarmag.com/PDF/LiDARNewsMagazine_Gupta-SmartGISand3D_Vol5No4.pdf

Dongxia Zhang received her M.S. degree in electrical engineering from the Taiyuan University of Technology, Taiyuan, Shanxi, China, in 1992 and her Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 1999. From 1992 to 1995, she was a Lecturer with Taiyuan University of Technology. Since 1999, she has been working at China Electric Power Research Institute. She is the co-author of four books, and more than 40 articles. Her research interests include power system analysis and planning, big data and AI applications in power systems. She is an Associate Editor of *Proceedings of the CSEE*.

Xiaoqing Han received her B.S., M.S. and Ph.D. degrees from the College of Electrical and Power Engineering, Taiyuan University of Technology, Taiyuan, China. She is presently working as a Professor at the Taiyuan University of Technology. Her current research interests include power system simulation, stability analysis, and the integration of renewable energy sources.

Chunyu Deng received his M.S. degree in computer engineering from the North China Electric Power University, Beijing, China, in 2008. Currently, he is studying for a doctorate in information security engineering. His research interests include power informatization, big data and artificial intelligence applications in power systems. He is currently the Director of the Artificial Intelligence Innovation Research Center of China Electric Power Research Institute.