IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 7 July 2022, accepted 26 July 2022, date of publication 29 July 2022, date of current version 8 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3195053

SURVEY

A Survey on Deep Learning Role in Distribution Automation System: A New Collaborative Learning-to-Learning (L2L) Concept

MINA JAFARI[1], ABDOLLAH KAVOUSI-FARD[1], (Senior Member, IEEE), MORTEZA DABBAGHJAMANESH[2], (Senior Member, IEEE), AND MAZAHER KARIMI[3], (Senior Member, IEEE)

[1]Department of Electrical and Electronics Engineering, Shiraz University of Technology, Shiraz 71557-13876, Iran
[2]Electric Reliability Council of Texas, Austin, TX 76574, USA
[3]School of Technology and Innovations, University of Vaasa, 65200 Vaasa, Finland

Corresponding authors: Mazaher Karimi (mazaher.karimi@uwasa.fi) and Abdollah Kavousi-Fard (kavousi@sutech.ac.ir)

This work was supported by the University of Vaasa under the Profi4/WP2 Project by the Academy of Finland and Business Finland under Grant 27081089141.

ABSTRACT This paper focuses on a powerful and comprehensive overview of Deep Learning (DL) techniques on Distribution Automation System (DAS) applications to provide a complete viewpoint of modern power systems. DAS is a crucial approach to increasing the reliability, quality, and management of distribution networks. Due to the importance of development and sustainable security of DAS, the use of DL data-driven technology has grown significantly. DL techniques have blossomed rapidly, and have been widely applied in several fields of distribution systems. DL techniques are suitable for dynamic, decision-making, and uncertain environments such as DAS. This survey has provided a comprehensive review of the existing research into DL techniques on DAS applications, including fault detection and classification, load and energy forecasting, demand response, energy market forecasting, cyber security, network reconfiguration, and voltage control. Comparative results based on evaluation criteria are also addressed in this manuscript. According to the discussion and results of studies, the use and development of hybrid methods of DL with other methods to enhance and optimize the configuration of the techniques are highlighted. In all matters, hybrid structures accomplish better than single methods as hybrid approaches hold the benefit of several methods to construct a precise performance. Due to this, a new smart technique called Learning-to-learning (L2L) based DL is proposed that can enhance and improve the efficiency, reliability, and security of DAS. The proposed model follows several stages that link different DL algorithms to solve modern power system problems. To show the effectiveness and merit of the L2L based on the proposed framework, it has been tested on a modified reconfigurable IEEE 32 test system. This method has been implemented on several DAS applications that the results prove the decline of mean square errors by approximately 12% compared to conventional LSTM and GRU methods in terms of prediction fields.

INDEX TERMS Cyber security, distribution automation system, deep learning, learning-2-learning.

ABBREVIATIONS

AAE	Adversarial auto-encoder.
Adaboost	Adaptive boosting.
AE	Auto-encoder.
AEN	Auto-encoder network.
AI	Artificial intelligence.
CDBN	Conditional deep belief network.
CNN	Convolutional neural network.
DAE	Denoising auto encoder.
DAS	Distribution automation system.
DBM	Deep Boltzmann machine.
DBN	Deep belief network.

The associate editor coordinating the review of this manuscript and approving it for publication was Yu-Da Lin.

DBSAE	Deep belief Stacked auto encoder.
DDPG	Deep deterministic policy gradient.
DL	Deep learning.
DNN	Deep neural network.
DNR	Dynamic distribution network reconfiguration.
DOS	Denial of service.
DQN	Deep Q net.
DR	Demand response.
DRBM	Deep restricted Boltzmann machine.
DRL	Deep reinforcement learning.
DRN	Deep residual network.
EDBN	Ensemble DBN.
EMD	Empirical mode decomposition.
ESNC	Echo state network classification.
EV	Electric vehicle.
FCBRM	Factored restricted Boltzmann machines.
FDIA	False data injection attack.
GAN	Generative adversarial network.
GBM	Gradient boosting Model.
GRU	Gated recurrent unit.
HVAC	Heating, ventilation, and air conditioning.
IMF	Intrinsic mode function.
LSPI	Least square policy iteration.
LSTM	Long short term memory.
LTLF	Long-term forecasting.
MAE	Mean absolute error.
MAPE	Mean absolute percentage error.
MBE	Mean bias error.
MDP	Markov decision process.
ML	Machine learning.
MLP	Multi-layer perceptron.
MLR	multiple linear regression.
MRE	Mean relative error.
MSE	Mean square errors.
MTLF	Mid-term forecasting.
NN	Neural network.
OPF	Optimal power flow.
PEV	Plug-in electric vehicle.
QR	Q-learning reinforcement.
RF	Random Forest.
RL	Reinforcement learning.
RBM	Restricted Boltzmann machine.
RMSE	Root mean square error.
RNN	Recurrent neural network.
SAE	Stacked auto encoder.
SC	Smart city.
SCADA	Supervisory control and data acquisition.
SDA	Stacked denoising auto encoder.
SDE	Standard deviation error.
SG	Smart grid.
SSAE	Stacked sparse auto encoder.
STLF	Short-term forecasting.
SVR	Support vector regression.
SVM	Support vector machine.
S2S	Sequence to Sequence.
TLN	Time LeNet.
VCC	Volt-VAR control.
WPD	Wavelet packet decomposition.

I. INTRODUCTION

Distribution Automation System (DAS) is an intelligent system that enables electric power utilities to remotely monitor, coordinate and operate distribution components in a real-time mode [1]. The main objective of DAS can be summarized as follows: improve voltage control, accurate load, and energy forecasting, improve the reliability and security of the system, accurate planning and operation of data, and improve fault detection/restoration and reconfiguration [2]. Generally, DAS is based on a technology that gathers and analyzes data for making control decisions according to the distribution power system model and then executes the suitable control decisions to achieve the desired result [3]. The traditional models for analysis, control, and decision-making of DAS are mostly dependent on the physical model of the distribution system and mathematical calculations. These methods did not achieve satisfactory outputs, with regards to reliable and accurate information of distribution network or customers is not available.

Artificial intelligence (AI) algorithms such as statistical methods, optimization theory, neural networks (NNs), and machine learning (ML) are examples of state-of-the-art technologies, which have been vastly applied to the complex problems of the power system. ML has emerged in the power area in the recent two decades that has the ability to learn from a huge amount of historical data and to make rapid decisions without human intervention. ML includes several algorithms that have been successfully employed for different disciplines such as classification, recognition, regression, prediction, and so on. DL is a subset of ML that applies cascaded layers to extract several features automatically of raw data. DL techniques have blossomed rapidly, and have been widely used in several fields. DL algorithms can be categorized, i.e., supervised, semi-supervised and unsupervised. In addition, Reinforcement Learning (RL) or Deep RL (DRL) is another classification of the DL algorithm [4]. Fig. 1 represents the different techniques of DL that are available in the literature.

The mathematical analysis and details of DL algorithms can be adapted from several papers like Deep Neural Networks (DNN) [5], Convolutional Neural Networks (CNN) [6], Recurrent Neural Networks (RNN) [7], Long Short Term Memory (LSTM) [8] and Gated Recurrent Units (GRU), Auto-Encoders (AE), Restricted Boltzmann Machines (RBM), Generative Adversarial Networks (GAN) and RL [9]. The use of DL has received a great deal of attention in recent years due to the unique features such as a robust, universal, scalable, and scalable learning approach that contributes to the sustainable development and security of modern power systems. DL algorithms have been proven as a worthy performance to solve the complex problems in the power system in several studies such as [10]. Hence, DL, RL, and DRL seem to be advanced tools to overcome the

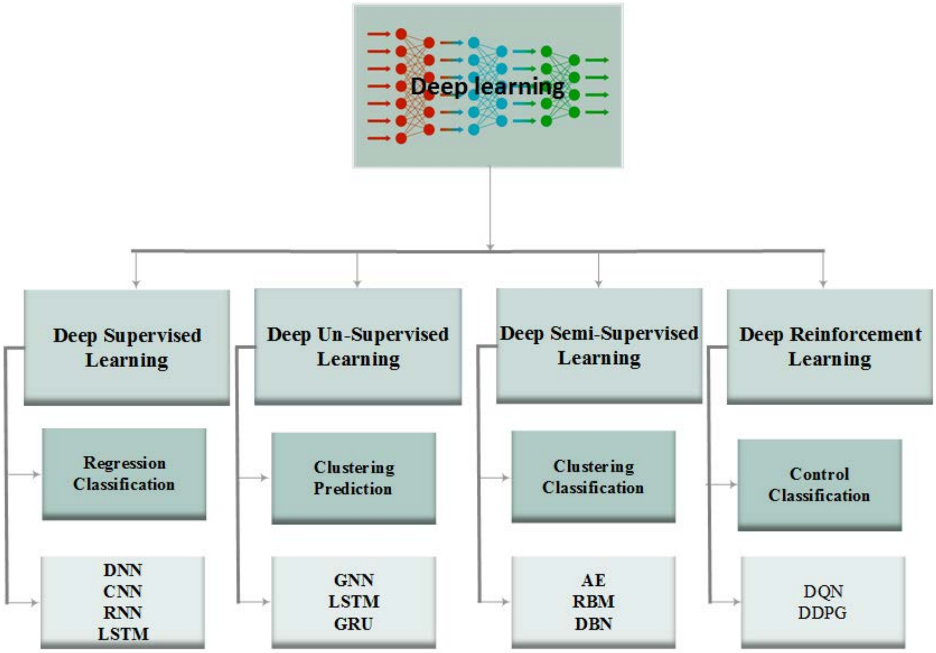

FIGURE 1. Classification of DL methods available in the literature.

problems of uncertainty and computational complexity based on a huge amount of data from the DAS. The manuscript presents a comprehensive overview of the DL techniques on the DAS, which considers several application areas. These applications include fault detection and classification, load and energy forecasting, electricity market, cyber security, network reconfiguration and restoration, voltage control and demand response.

The mentioned applications are categorized into two categories. The first category includes applications based on classification and prediction that have been studied by deep supervised, semi-supervised and unsupervised learning methods. The second is used when the problem is based on decision-making (RL method) that leads to the optimization of the desired objective. Classification of DL methods employed in this study illustrates in Fig. 1.

In addition to the reviewed DAS application in this manuscript, the L2L platform-based DL is suggested for solving the problems of DAS. Section II presents DL techniques for applications of DAS that are divided into subchapters on cyber security, voltage control, network reconfiguration and restoration, etc. Section III presents the proposed method structure for future research. The general evaluation of DL methods is conferred in section VI as discussions. Finally, the conclusions are expressed in section IV. Also, a summary of DL methodologies in the literature is presented at the end of this section.

II. APPLICATIONS OF DL IN DISTRIBUTION AUTOMATION SYSTEM

This section reviews the DL algorithms in applications of DAS. The existing literature is divided into sections such as load forecasting, energy forecasting, fault detection and classification, energy market forecasting, network reconfiguration and restoration, demand response, cyber security and voltage control. As mentioned above, these applications are grouped into two categories according to their field of work. Fig. 2 depicts the DAS applications reviewed in this section. Also, the common evaluation indexes for comparing the efficiency of the DL techniques are investigated.

A. LOAD FORECASTING

Load forecasting is one of the most important influencing factors in the planning, operation and control of modern power systems. The high penetration of distributed energy resources into the existing grid increases the uncertainty of operations and planning of SG [11]. Therefore, the correct forecasting of load at different levels is beneficial for the DAS economically and saving electricity. Load forecasting depends on several factors that can change the load consumption pattern. The time horizon factor is one of them that includes Short-term forecasting (STLF), Mid-term forecasting (MTLF) and Long-term forecasting (LTLF). STLF has been used in different fields, such as economic load dispatch, real-time control, energy transfer scheduling, and demand response [12]. MTLF and LTLF can be applied for planning the power plants and represent the dynamics of the power system [13]. Several intelligent algorithms are used for load forecasting based on small datasets [14]. Hence, DL algorithms considering the larger volumes of datasets extracted by smart meters are efficient for load forecasting in the DAS [9].

A summary of the DL methods used in SG for load forecasting is presented in this paper. In order to plan and operate of power system, a DNN for STLF is presented in [15]. Then, the probability density of load consumption has been forecasted based on the DNN method combined with quantile

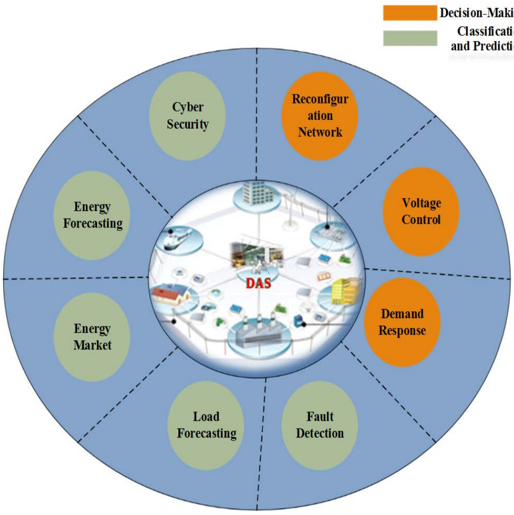

FIGURE 2. Categories of DAS applications for DL methods.

regression. Lastly, the results of the suggested method are compared with ML tools such as random forest and gradient boosting machine. Hence, there exist different types of load forecasting papers that apply DL application in the DAS to STLF. In [16], the application of Feed-forward DNN and Recurrent-DNN models by use of STLF data to compare are studied. Synthesis of CNN and K-Means algorithms for STLF with improved scalability is suggested in [17] by using K-means algorithm, the large volume data set is clustered in to proper subsets. Then subsets are trained as input to CNN. The experimental results of the proposed method are shown its effectiveness. The DNN combined with CNN for STLF in a north China city is used [18]. CNN approach is used to learn deep features from the historical dataset. The variation of historical load dataset is modeled using LSTM based RNN. Likewise, the learned dataset is used to predict load via dense layers. The proposed method is quite flexible and efficient that can be applied to other predictions. A DBN compound with parametric Copula models for forecasting hourly load is suggested in [19]. Load data in an urban area in Texas is utilized for the experimental validation. Proposed methods are compared with ANN, SVR, and ELM. By mean MAPE and RSME, the experimental results have confirmed that the proposed model is an effective method. In [20] is proposed an approach using DBN that is made from multiple layers of RBMs for STLF in the Macedonian for validation of the proposed method. Here, the layer-by-layer unsupervised training method is controlled by fine-tuning the parameters by using a supervised back-propagation training method. DRNN–GRU model for STLF and MTLF by using consumption data of load building is presented [21]. This method is evaluated by using several factors such as the root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). This model is compared to that of DRNN-LSTM, RNN, MLP, ARIMA, SVM, and MLR. The DRNN-GRU model can be achieved better performance for estimating the load demand data. Authors in [22] are represented empirical mode decomposition (EMD) based DL which incorporates the EMD approach with the LSTM model to forecast the electric load demand for a time interval time. The EMD technique is analyzed the load data time series signal into various intrinsic mode functions (IMFs) and residue. Afterward, the LSTM model is trained by using IMF as input data. Finally, to determine the combined output for electricity demand, the prediction outcomes of all IMFs are incorporated together. In [23] is proposed Load demand Forecasting using EMD method composed with DBN containing two RBMs of energy in Australia. Furthermore, the performance of the proposed approach is evaluated by comparing the prediction results with SVR, ANN, DBN, RF, EDBN, EMD–SVR, EMD-SLFN, and EMD-RF models. A residential load forecasting by using CNN is suggested in [24]. In order to decrease the MAE, the suggested model is combined with CNN. The CNN method also is compared with other techniques like SVM and ANN. Stacked denoising auto-encoders (SDAs) model is proposed for electricity load forecasting [25]. The output of SDAs data is used for the training process of SVR model as input. In [26] is proposed two LSTM methods based on hourly and minute ahead for load demand prediction. Here, the output results have confirmed that a standard LSTM method cannot accurately predict while the LSTM-based Sequence to Sequence (S2S) method achieved accurate prediction. Also, authors in [27] by using historical data are suggested CNN method for load demand predictions. Output results from CNN are compared with LSTM- S2S, FCRBM, shallow ANN and SVM. Experiment results are shown that CNN has a better performance compared to other methodologies. Due to inconsistent consumption patterns in utility customers, LSTM–RNN approach for load demand forecasting is studied in [28]. An innovative pooling-based deep-RNN for load forecasting in Ireland residential is proposed in [29]. In [30] to amend the accuracy of load forecasting in short term, a hybrid technique that includes DBN, LSTM and SAE is proposed. Table 1 presents a detailed overview of DL methods in load forecasting. The RMSE, MAE and MAPE are useful evaluation indexes that are applied for comparing the results.

B. ENERGY FORECASTING

Renewable energy sources, often called clean energy play a significant role in power distribution systems. However, renewable energy systems often have uncertain characteristics, which lead to uncertainties in energy power systems [11]. Power systems are complicated artificial systems. With the evolution of the smart grid, high penetration of wind, solar power and customer participation have led systems to operate in more complex and uncertain environments. Traditional power system analysis and management decision-making are related to physical modeling and numerical computations. The conventional methods find difficulty in handling uncertain subjects so that they cannot satisfy the essentials of the future evolution of smart grids. On the other hand,

TABLE 1. Summary of some selected DL methods and evaluation indexes in load forecasting.

Ref.	Type of method (Comparative methods)	Type of application	Evaluation index				
			RMSE	MAE	MAPE	MRE	MBE
15	DNN(RF, MLP)	Load Forecasting	✔		✔		
16	R-DNN, FF-DNN(DNN)		✔	✔	✔		
17	CNN, K-Means(SVR, CNN)		✔		✔		
18	CNN, LSTM–RNN, (CNN-RNN,DNN, LR, SVR)			✔	✔		
19	Copula- DBN (NN, SVR, ELM, DBN)		✔		✔		
20	DBN-RBM (traditional model)				✔		
21	DRNN-GRU (LSTM, RNN, MLP)		✔	✔	✔		
22	LSTM (RNN-EMD)		✔				
23	DBN –RBS (SVR, ANN, RF)		✔	✔			
24	CNN (SVM , ANN)				✔		
25	SAD (SVR, ANN)				✔		
26	LSTM		✔				
27	CNN (LSTM, ANN)		✔				
28	LSTM-RNN (BPNN, KNN, ELM)		✔				
29	DRNN (ARIMA, SVR)		✔	✔			
30	DBN-LSTM, SAE		✔	✔	✔		
31	CNN (AE, DNN, Shallow)			✔	✔		

the extensive presence of advanced measurement infrastructure (AMI) and monitoring/management systems generate massive data and equip the foundation of data for model training in deep learning applications. Hence, DL, RL, and DRL appear to be some of the most robust technologies for the future development and success of the modern power grid. In fact, uncertainties of solar and wind energy bring many challenges to power systems. DL is a powerful tool to enhance solar, and wind generation forecast accuracy based on large datasets. Also, RL can help sequential decision-making under uncertainty. The RL algorithm functions with only restricted knowledge of the environment and with restricted feedback on the quality of the decisions. Scenario generation can assist to model the uncertainties and variations in renewables generation, and it is a basic means for decision-making in power grids with high penetration of renewables.

A data-driven approach based on renewable scenario generation using GANs is proposed in [32]. This method compared with probabilistic methods is data-driven and can be captured renewable energy production patterns in both temporal and spatial measurements for a large number of correlated resources. In [33], the authors used the DBN method for wind and PV power prediction. This method due to capturing uncertainties in the energy time series has led to accurate prediction. In [34], the convolutional graph AE method is proposed for solar irradiance prediction. This method is applied in applications where the temporal and spatial correlations are vital points. Based on the results, the proposed method had the lowest RMSE compared to LSTM. In [23], the authors proposed a new interval probability distribution learning method to increase the accuracy of planning and scheduling of wind energy generation. The proposed approach is presented for learning temporal traits from the time-series data to handle the uncertainty of renewable energy. Therefore, the accurate prediction of renewable energy for planning and operation of the power distribution system is an important way to face this problem. In recent years, DL techniques have been proposed for wind power and PV systems forecasting. A hybrid model that includes wavelet transform, DBN and spine-QR for short-term wind speed forecasting in China is proposed in [35]. The wavelet transform is applied to analyze the wind speed time datasets. Then, the invariant features of each dataset are obtained by DBN. Finally, the uncertainties in wind speed are considered by the QR technique. In [36] is suggested a wind power forecasting technique based on wavelet transform and CNN. The nonlinear features from each dataset are learned by the CNN method. The proposed model is compared with Back-Propagation ANN and SVM of ANN architectures. The prediction results demonstrate the supremacy of the stated method to predict the probabilistic wind power. Empirical-WT, LSTM and Elman neural network techniques are combined to wind speed forecasting. The empirical-WT is implemented to analyze the wind speed time dataset into numerous datasets. The LSTM and ENN techniques are applied to forecast low/high frequency datasets [37]. In [38] DNN method that includes LSTM and CNN is proposed for deterministic short term wind power forecasting in northeast China. In order to best perform probabilistic forecasting, the results of the deterministic model are evaluated. In [39] a DBN method for wind forecasting in short term is proposed. Then, the k-means clustering algorithm is used to face the uncertainty in wind power. In [40] a technique that combines AEs and DBM is presented for short term wind power forecasting. The outcome results in MAE and RMSE terms are derived better performance compares to other techniques. The DBN, SAE and DSAE algorithms are proposed for wind speed forecasting in Brazil [41]. Similarly, in [42] rough-NNs with SAE and SDAE methods are combined that called rough-SAE and rough-DAE for forecasting the accuracy of the wind power. Output results are shown improved performance in comparison to SAE and SDAE approaches. In [43] LSTM and RNN methods are studied for forecasting PV power in Egypt. In order to achieve the best forecasting, different LSTM approaches are tested. In [44] several approaches such as MLP, LSTM, DBN and AEs are suggested for solar power forecasting. The performance of the methods is evaluated on a dataset containing solar power measurements in Germany. The DRNN–LSTM model for PV power forecasting is proposed [45]. The experiment results are exhibited the best performance compared to MLP and SVM methods. In [46] are proposed a DNN model called PVPNet to PV power forecasting that is used the solar irradiance dataset as input variables. In [47] is applied DBN to short term power PV prediction. In [48] a technique is proposed for PV power forecasting, which combines wavelet transform and DCNN. The method is tested on PV farm data-sets in Belgium. That is derived better results in terms of MAPE, RMSE, and MAE compared to other methods. Similarly, the CNN model is presented for forecasting PV power on a short time horizon [49]. A deep learning-based hybrid approach for short-term solar PV power prediction is proposed in [50]. CNN, LSTM, and ANN are combined for mapping between surface irradiance measurement and the sky image information. The results demonstrate the effectiveness of the proposed approach based on DL compared with current methods (such as CNN, LSTM, and ANN). In [51], the authors presented a hybrid deep convolutional CNN-LSTM technique to predict horizontal irradiance. Dataset was collected from three solar stations in the east of the United States. In [47], [52] the CNN method to forecast PV power is proposed that combined with SVM and LSTM methods. The results of the evaluation criteria are shown that the proposed model could perform better than other methods. A hybrid DL technique combining wavelet packet decomposition (WPD) and LSTM is suggested for the prediction of PV with high accuracy in comparison with LSTM, GRU, RNN and MLP techniques in the presence of power curves data of PV [53]. A GAN-based model for renewable power forecasting in microgrids is proposed in [54]. The concept of GANs is utilized to forecast the output power of tidal and photovoltaic units. The results showed the high performance of the proposed methodology compared with SVR and artificial NNs. In [55], a wind power prediction technique applying ensemble learning and transfer learning based on DNN is proposed. Ensemble learning is applied to facilitate robust decisions on data that are unseen. Also, transfer learning is used to facilitate fast learning on wind data. The performance indexes were RMSE, MAE, and MAPE. Based on the results, the proposed method had the highest performance in predicting the wind power variable. These studies are provided a detailed overview of DL methods in energy forecasting that are presented in Table 2.

C. ENERGY MARKET FORECASTING

Energy markets play an essential role in the power system as they make SG cost efficient. Energy markets benefit from load and price forecasts. Hence, accurate electricity price forecasting would help the market participants in order to barrier against future price changes and maximize their profits. There are several methods for price forecasting in recent studies. DR methods are superior to other methods with regard to managing big price data in the network. A hybrid deep neural network model which synthesizes the CNN and LSTM for electricity price forecasting is presented in [56].

TABLE 2. Summary of some selected DL methods and evaluation indexes in energy forecasting.

Ref.	Type of method (Comparative methods)	Type of application	Evaluation index				
			RMSE	MAE	MAPE	MRE	MBE
35	DBN(ARMA, BPNN, MWNN)	Energy Forecasting		✔			
36	CNN-DCN,WT , DCNN (BPNNWT,SVM)		✔				
37	CNN, LSTM (ARIMA,SVM,BP,RBF, ENN, ELM)		✔	✔	✔		
38	LSTM , CNN		✔	✔	✔		
39	DBN		✔	✔	✔		
40	SAE, DBM		✔	✔			
41	DBN, SAE, DSAE		✔	✔			
42	RSDAE (PR, FFNN, TDNN,NARNN)		✔	✔			
43	LSTM-RNN (MNL, BRT, NN)		✔				
44	MLP, LSTM, DBN , SAE		✔	✔			
45	DRNN–LSTM (MLP,SVM)		✔	✔	✔		
46	Deep-CNN (DT, RF, SVM,MLP)		✔	✔			
47	DBN (SAE-DBM)		✔		✔		
48	WT, QR, DCNN			✔	✔		
49	VMD–CNN		✔	✔	✔		
50	CNN-LSTM-ANN (hybrid CNN-LSTM,CNN-ANN)		✔	✔			
51	CNN (SVM, LSTM)				✔	✔	
52	CNN (LSTM, SVM, RBF, BPNN)				✔		
53	Hybrid LSTM- WPD (LSTM, GRU, RNN, MLP)		✔		✔		✔

TABLE 3. Summary of some selected DL methods and evaluation indexes in the energy market.

Ref.	Type of method (Comparative methods)	Type of application	Evaluation index				
			RMSE	MAE	MAPE	MRE	MBE
56	CNN- LSTM (ANN, SVM, MLP, CNN , LSTM)	Energy Market Forecasting	✔	✔			
57	GRU-LSTM (RNN)			✔			
58	LSTM-DNN,GRU-DNN (MLP, CNN)		✔	✔			
59	SDA (DNN)				✔		
60	LSTM (DNN)		✔				
61	DRNN (SVM)			✔	✔		
62	LSTM				✔		

The result of the study has proven that the proposed model outperforms compared to other ML methods such as ANN, DT, MLP, SVM, MLP, CNN and LSTM. In [57] deep neural network model which combines the GRU and LSTM to forecast electricity price is proposed and compared with RNNs. Similarly, several methodologies that include LSTM-DNN, GRU-DNN, CNN and MLP for electricity price prediction are proposed in [58]. The hybrid algorithms (LSTM-DNN, GRU-DNN) are comparatively better than CNN and MLP. Meanwhile, the use of SDA and DNN models for this application is studied in [59]. In [58], four different DL algorithms such as DNN, LSTM, CNN and GRU are proposed to forecast electricity prices. DL models are compared with several ML approaches, the numerical results confirmed that the DNN model outperformed the other models. The deep LSTM model is proposed for load and price forecasting in [60]. The performance of the proposed model is confirmed by using real market data. In [61], the DRNN approach for forecasting day-ahead electricity price is studied and compared with the single and hybrid SVR. The outcome of the proposed DRNN method is more effective than others. In [62] is presented a novel day-ahead prices forecasting using a combined model which integrates LSTM with the attention mechanism for the electricity market. The main feature of the proposed model is that each sample considers a context vector as the cluster center to achieve a better prediction compared with similar LSTM models. Table 3 represents reviewed papers for DL methods used in the electricity market papers.

D. CYBER SECURITY

The main aim of cyber security is to protect cyber infrastructures from cyber-attacks. Cyber-attack against power systems infrastructures as challenging problem impacts on reliability, quality and security of smart systems. Attacks can manipulate the measurement data without the false change being detected. Early Detection of an attack will secure the power system. Several ML algorithms based on detection exist against cyber-attack. In [63], the CNN algorithm is proposed to detect replay attacks. The proposed model has been compared with other ML and DL models such as MLP, deep MLP (DMLP), deep residual network (DRN), time LeNet (TLN) model and echo state network classification (ESNC) to evaluate the high detection accuracy. Applications of DL for cyber-attack detection have been widely utilized in SCs areas. In [64], a conditional deep belief network (CDBN) based on a distributed DL algorithm has been suggested for the detection of thief electric in SC. The high detecting accuracy has demonstrated the effectiveness of the proposed method. A stacked deep polynomial network for intrusion detection is applied in [65] that can classify datasets into normal and attack data. DBNs are studied in [66], [67] for the same task. The results show the efficient performance of DBN in attack detection. In [68], hybrid models which combined DBN, MLP and RBM are applied for Denial of service attack (DOS attacks) detection for EV in SC. The SAE approach to detect the manipulated data in the SG is suggested in [69]. Similarly, for prediction and detection of the power system security weak spots, SAE is suggested [70] that output results have been proven the model has a simple implementation with low training time. SAE has yielded an average prediction accuracy of 95.78 % in the real system in china. A hybrid model which combined CNN and LSTM algorithms are proposed in [71], where the model was used for electricity theft detection. CNN is used for SG data extraction and classification. A combination of CNN, LSTM and SAE structures is presented in [72] for a similar purpose. Attacks are recognized as a serious threat to the supervisory control and data acquisition (SCADA) system. To address this issue, the CDBN algorithm is suggested in order to identify the False Data Injection Attacks (FDIAs) in SG [73]. It is shown that the proposed algorithm detects FDI attacks with detection efficient evaluation criteria (e.g. detection accuracy above 93%), compared to the ANN and SVM methods. To secure the SG, the wide and deep CNN model is suggested that used for electricity theft detection [74]. The wide component is used for learning and memorization of the global knowledge while the deep CNN component classified the non-periodicity and the periodicity of electricity consumption data. Similarly, in [75], the authors proposed an ensemble CNN model for electricity theft detection in SG.The output results are proven the best performance compared with other types of methods such as SVM, gradient boosting model (GBM), Random Forest, and DCNN. In [76], two different types of attacks based on FDIA have been detected by the use of the MLP method. In [77], authors suggested a semi-supervised learning approach based on adversarial AE (AAE) for detecting FDIAs distribution systems. The results illustrated that the proposed model achieved a high accuracy compared to AE and SVM algorithms. In [78], a deep-learning-based CNN algorithm is proposed for network intrusion detection for the SCADA

TABLE 4. Summary of some selected DL methods and evaluation indexes in cyber security.

Ref	Type of method	Type of attack	Comparative methods	Type of application	Evaluation index				
					Accuracy	FPR	Sensitivity	Specificity	precision
63	CNN	Replay attack	MLP,DMLP, DRN , TLN, ESNC	Cyber Security	✔	✔	✔	✔	✔
68	DBN	DDoS, DoS, Probe, U2R, L2R	DNN		✔	✔			✔
70	SAE	Cyber-attack	Shallow model		✔				
71	CNN-LSTM	Cyber-attack	SVM		✔				✔
72	Deep CNN LSTM-SAE	Cyber-attack	Shallow model		✔		✔		✔
73	CDBN	FDIAs	ANN, SVM		✔	✔	✔	✔	
74	CNN	Cyber-attack	SVM, RF, CNN		✔				
75	CNN	Cyber-attack	SVM, Random Forest, DCNN		✔		✔		✔
76	MLP	FDIA	SVM, DBM, RBN		✔		✔		✔
77	AAE	FDIA	SVM-AE		✔		✔		✔
78	CNN	Network based attacks	-		✔				✔
79	CNN	Cyber-attack	-		✔	✔	✔	✔	✔
80	2-D CNN	FDIA	SVM, 1-D CNN, LR		✔		✔		✔
81	RNN	Cyber-attack	SVM			✔	✔		

system. The results are shown high accuracy for attack detection in real-world SCADA systems that detection accuracy was 99.84%. In [80], since the electricity data structure is a kind of time-series data, two-dimensional CNN is developed for electricity theft detection. In [81] a high precision RNN model named LSTM–UNet–Adaboost to detect electricity theft is proposed. To improve the performance of the theft detection, this method applies the DL technique and ensemble learning. In comparison to SVM, the accuracy index has been improved by 39.6%. In [82], a novel Cyber security method based on the GAN structure is presented. The attacker to the power grid follows two aims. The first objective is to be hidden from the system defender, and the second purpose is to earn profit through its FDI into the system. The results show the proposed model detects FDI very well and with high accuracy. These studies are summarized in Table 4. The accuracy, FPR, sensitivity, specificity and precision are evaluation criteria that are commonly used to evaluate the algorithms in cyber security fields as shown in Table 5.

TABLE 5. Cyber security evaluation indexes.

Evaluation index	Description
$\text{Accuracy}=\frac{TP+TN}{TP+FN+FP+TN}$	
$\text{FPR}=\frac{FP}{FP+TN}$	FPR is false positive rate
$\text{Sensitivity}=\frac{TP}{TP+FN}$	TP is true positives TN is true negatives FP is false positives FN is false negatives
$\text{Specicity}=\frac{TN}{TN+FP}$	
$\text{Precision}=\frac{TP}{TP+FP}$	

E. FAULTS DETECTION AND ISOLATION

Faults in the distribution system lead to power outages, so cause huge losses that can result from short circuits, overloading, human mistake, and so on. Fault detection and isolation are other functions of DAS that lead to enhancing reliability as well as efficiency and quality of electric distribution systems. Fault detection is the procedure of analyzing historical data to find a fault in reliable power systems [83]. In [84], the authors suggested a hybrid algorithm that combines LSTM networks and SVM for line trip fault prediction in power transmission and distribution. LSTM networks are applied for training and mining the temporal features of data. Then, SVM is used for feature classification in order to achieve prediction results.

An approach based on SSAE is proposed for fault detection [85]. To improve accuracy the SSAE network is combined with SVM and principal component analysis. The DL approaches have been studied for fault detection in the secondary distribution network [86]. The approaches include GRU, RNN and LSTM. The simulation results are

TABLE 6. Summary of some selected DL methods and evaluation indexes in fault detection.

Ref.	Type of method	Comparative methods	Accuracy (%)	Type of application
84	LSTM-SVM	BPNNs, SAEs, RNNs , SVM	97.7	Fault Detection
85	SSAE	Decision Tree ,SVM,BPNN , DBN	91.8	
86	RNN –LSTM-GRU	MLP ,DMLP DRN ,TLN, ESNC	94	
87	CNN	-	Over 85	
88	DBSAE	KNN, SVM, BPNN	95	
89	SAE - DBN	Shallow model	accurate	
90	SAE	BP	71.3	
91	DNN	SVM, GBM, DCNN	accurate	
92	ACNN	SVM	above99	
83	CNNs	-	90	
93	CNN	SVM	above 98.5	
94	CNN	SVM	above 99.5	

evaluated real-time measurements by the dataset from 2014 to 2020 that show the RNN accuracy of 94% and GRU and LSTM methods of 50%. In [87], [88], the authors suggested the adversarial CNNs which combined the GANs and CNNs for Fault detection and isolation and repair. A CNN algorithm is proposed for fault classification of the power systems. This algorithm has achieved an accuracy of over 85% for per phase and three phase testing. To detect the transformer fault, a deep belief SAE approach is suggested [89]. The results revealed better performance in comparison with KNN, SVM, BPNN. Similarly, the transformer fault is predicted using SAE and DBN algorithms to increase reliability and the stability of power systems [90], [91]. The SAE algorithm is proposed in [83] for the same task. The hidden layers in this algorithm have increased the accuracy of detection. A combined method with wavelet transform and DNNs which provided fault type, fault phase and fault location in microgrid systems is presented in [92]. The proposed method is indicated more accurate prediction results compared with conventional methods. Another fault identification method based on the DBN algorithm is applied to underground cables which are extensively used in distribution systems [93]. An adaptive CNN base on fault diagnosis for distribution network fault location is proposed in [94]. The advantages of the proposed method are low computation time and the high accuracy/speed of fault line selection. Also, in [95] authors addressed CNNs algorithm in order to get the fault type/location for a distribution system. The simulation results have proven an accurate performance of CNN in comparison with other techniques, such as SVM. In [96], a continuous wavelet transform and CNN is proposed for faulty feeder detection in power distribution systems. Similarly, for the same work, the CNN method has been suggested in [97]. In [98], a fault recognition method of the voltage sampling module of distribution terminals based on GAN and CNN is presented. The GAN model is used to generate the pattern and learn the developed samples then these samples are used to train the CNN method. This combination method can significantly improve the accuracy of fault detection. In [99], the authors proposed a novel transfer learning technique to assess multiple faults which are not trained. Validation of the proposed method was performed by comparing the dynamic security assessment model. Transfer learning could assess fault with 97.27% of accuracy. In [100], a fault detection approach using GANs for small-sample WT based on the data collected from the wind farm SCADA system in Northern China is presented.

These studies are provided a detailed overview of DL methods in the detection and classification of fault that are presented in Table 6.

F. NETWORK RECONFIGURATION AND RESTORATION

Dynamic distribution network reconfiguration (DNR) is an optimization decision-making process that changes the hourly status of remotely controllable switches to improve the performance of distribution systems. The DNR has represented great potential in enhancing several important aspects of the electricity network operation. It can be used for minimizing power losses [101] or the system costs [102], [103], improving the voltage profile [104], [105], increasing load balance [106] or the system's reliability [107], [108]. The DNR problem is a mixed-integer nonlinear problem that depends not only on uncertainties, regarding the active and reactive powers but also on the physical parameters of the system. The dynamic DNR problem can be classified into three groups: mixed-integer programming, heuristic algorithms, and dynamic programming methods. The first and second are based on model-based algorithms that may not be reliable due to the inaccuracy of the distribution network parameters. In addition, the computation time for model-based control algorithms increases due to network topology. Therefore, dynamic programming methods have been applied to solve the DNR problem. To bypass these issues, the RL method is used in [109], [110], where a batch-constrained algorithm is proposed to learn from the historical reconfiguration data collected by the distribution system operator without distribution network parameter information, which enables to avoiding of inaccuracies from the representation of the physical grid. The objective of the proposed algorithm is to minimize the system operational cost and loss. A real-time autonomous dynamic reconfiguration method to reduce the cost of power loss and switch action of the distribution network based on the DL algorithm is proposed in [111] that can achieve a reconfiguration solution in the order of milliseconds and has high robustness. A hybrid multi agent Q-learning algorithm to determine the changing switches for system restoration in a timely manner is proposed in [112]. A shipboard power system reconfiguration algorithm based on Q-learning is proposed in [113]. The paper obtained the best sequence of open/close switches to do a final configuration which takes the shortest amount of running time. A restoration algorithm using multi-agent Q-learning in order to find switching configurations is proposed in [114]. Similarly, for optimization of switching status and reducing power losses a Q-learning framework for distribution network restoration is proposed in [115]. A tabular Q-learning algorithm for network reconfiguration to opening/closing the switch status with the objective of minimizing power losses is presented [116].

G. VOLTAGE CONTROL

Volt-VAR control (VVC) is a critical application in distribution system automation to reduce network losses and improve voltage profiles. To remove dependency on inaccurate and incomplete network models and enhance resiliency against communication or controller failure, DRL methods are used in several papers and applied Q-learning for reactive power control [117]. A multi-agent Q-learning VVC framework is proposed in [118] in order to reduce the communication and computation burden of a central controller. In [119], the least square policy iteration (LSPI) algorithm is developed to control tap changer positions that are effective to diminish the voltage deviation. This algorithm was introduced as batch RL and adopted to handle scalability. By using the LSPI iteration, an approximation of the Q-function is constructed. The constrained soft actor-critic algorithm to solve the VVC problem is studied [120]. In this paper violation of the voltage is considered as a constraint, also power loss and switching cost are rewards in the suggested algorithm. In [121], the quadratic programming and deep Q net (DQN) agent are suggested to solve control problems in fast time-scale and slow time-scale. In [122], the multi-agent MDP is formulated for the VVC problem and proposed as a multi-agent deep Q learning method. In order to improve the learning efficiency, the action space is decomposed in each device method. The consensus multi-agent DRL algorithm to solve the VVC problem by using the maximum entropy RL framework is proposed [123]. This problem is modeled as MDP. The results demonstrate the effective performance of the proposed framework. The CMDP is formulated for the VVC problem that the voltage violations are considered as constraints [124]. A constrained policy optimization algorithm is extracted to solve the MDP problem. In [125], the deep deterministic policy gradient (DDPG) is proposed to modify the voltage profile and diminish the constraint of PV generation. An emotional DL programming controller for voltage control of the distribution system is proposed [126] that includes an emotional deep neural network and an artificial emotional Q-learning algorithm. Outcome results are proven the effectiveness of the proposed algorithm compared with the DNN and Q-learning algorithms.

H. DEMAND RESPONSE

Demand response (DR) is a power consumption variation that keeps a balance between power demand and generation. DR can improve the flexibility of the system through shift peak or valley of energy consumption in real time. DR programs are classified into two main categories, dispatchable or incentive based programs and nondispatchable or price-based programs which are divided into subgroups as shown in Fig. 3 [127]. DRL algorithm is an effective model to solve control problems due to merging consumer and consumption in the control loop. An incentive based real time DR algorithm for SG using RL and DNN is proposed [128]. The purpose of this algorithm is to aid suppliers to purchase energy from various customers in order to balance power variations and improve the validity of the grid. DNN is applied to predict unrevealed prices and energy demands. RL is used to acquire the optimal incentive cost for different customers by considering the profits of customers and suppliers. Similarly, in [129] a DRL technology is proposed to help decision-making in air conditioning and heating systems during DR to achieve an optimal control objective. In order to optimize the HVAC electricity consumption and minimize the total cost, a multi-agent RL is suggested in [130]. An optimal pricing schedule for a demand response scheme by RL is developed in [131]. The simulation results have shown the best efficacy of the proposed system. To solve the decision-making problem under uncertainty, RL is suggested that has modeled the energy consumption scheduling of several residential customers [132]. A novel DR manner in order to decrease the cost of charging/discharging plug-in electric vehicles (PEV) through a batch RL algorithm is proposed in [133]. A dynamic pricing DR approach is suggested in [134] that considers the service supplier's profit and customers' costs. RL is applied to solve the dynamic pricing problem due to demand and prices of power. For determining an optimal bidding strategy to purchase power from the grid, a multi-agent RL (Q-learning) algorithm is applied in [135]. In [136] a price-based DR scheme for industrial energy is studied that used the RL algorithm to optimize industrial energy. Simulation results have shown that the DR scheme could reduce energy costs and consider the balance between energy consumption and the consumer. In [137] several papers are analyzed that used RL for DR applications such as (HVAC, EV and energy storage) in SG.

III. ADVANTAGES AND DISADVANTAGES OF DEEP LEARNING TECHNIQUES

Table. 8 summarizes the advantages and disadvantages of various deep learning techniques in power knowledge investigation.

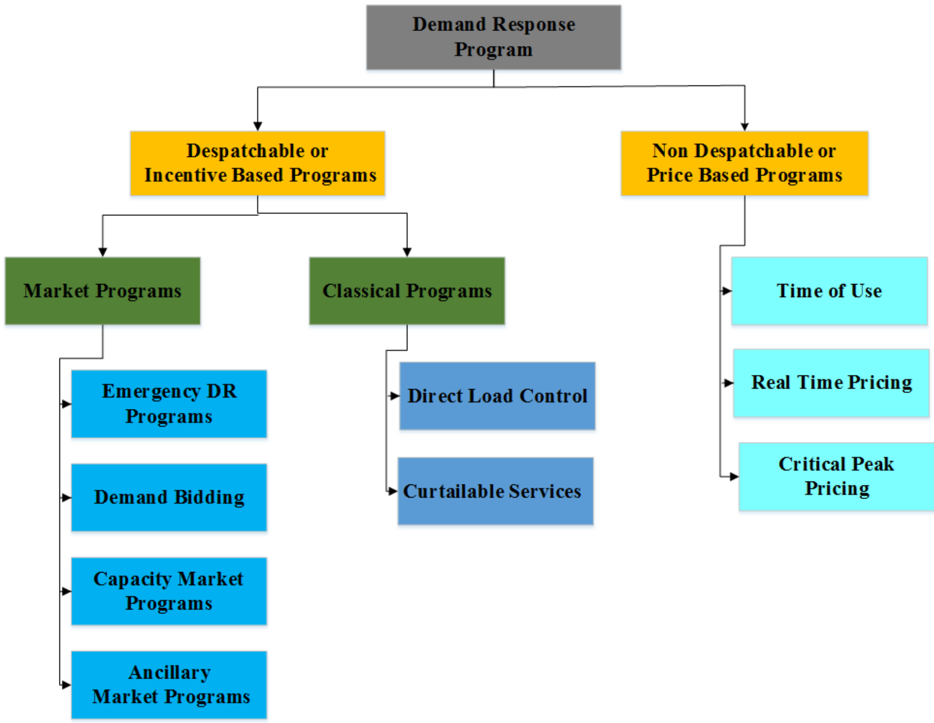

FIGURE 3. Type of demand response programs.

TABLE 7. Summary of literature reviews on distribution automation system applications.

Ref	Application	DL methods	Evaluation index	Objectives
[15-31]	Load Forecasting	CNN,DNN, LSTM,DBN,RBM,SAD,SAE	RMSE, MAE, MAPE	Improve prediction accuracy
[32-55]	Energy Forecasting	CNN, LSTM,DBN,AEN, SAE	RMSE, MAE, MAPE, MRE, MBE	Improve prediction accuracy
[56-62]	Electricity Market	DNN, LSTM, CNN , GRU,SDA	RMSE, MAE, MAPE	Increase consumer benefit, Minimize energy and operation cost
[63-82]	Cyber Security	CDBN,CNN,DBN,SAE-LSTM,AAE	Accuracy, FPR, Sensitivity, Specificity, precision	Security and detection of type of attack
[83-100]	Fault Detection	LSTM-CNN-DNN-SAE-GRU-DBN	Accuracy	Improve detection accuracy
[109-116]	Network Reconfiguration and Restoration	Q-learning, DRL, DQN, RL	Reward function	Reduce network losses, Minimize operation, Power loss and switching costs
[117-126]	Volt-VAR Control	DDPG, Q-learning, DQN	Reward function	Stability of energy and voltage control, Improve voltage profile
[128-137]	Demand Response	RL, DRL, Q-learning, DQL	Reward function	Load control, Improve energy consumption, Reduce energy costs

IV. DISCUSSION

This section presents conclusions and recommendations yanked from outcomes of investigations executed according to DL approaches in the DAS applications. As mentioned, Fig.2 demonstrates the categorization of DAS applications applied in the line with the purpose of the current study. According to overviewing results, it is noticed that LSTM and its hybrid techniques such as CNN-LSTM, LSTM-RNN methods have been mostly employed in the literature so far in the forecasting field. Since most studies utilize the RMSE evaluation index, therefore, it is the most used in DL techniques compared to the other evaluation indexes. In general, the contribution of hybrid methods to the various DL techniques used by the authors is greater than that of single methods and also hybrid methods have higher performance. Therefore, this domain is inclined to the usage of hybrid methods. From Table. 9, it can be asserted with certitude that DL techniques along with the RMSE lowest value hold the best performance compared with other techniques. Table.9 indicates output values for some studies developed with different DAS techniques and DL methods (studies no. 1 to 6 for load forecasting, studies no. 7 to 17 and 18 to 20 for energy and market forecasting, respectively and studies no. 21 to 27 for security), according to the table, most of the

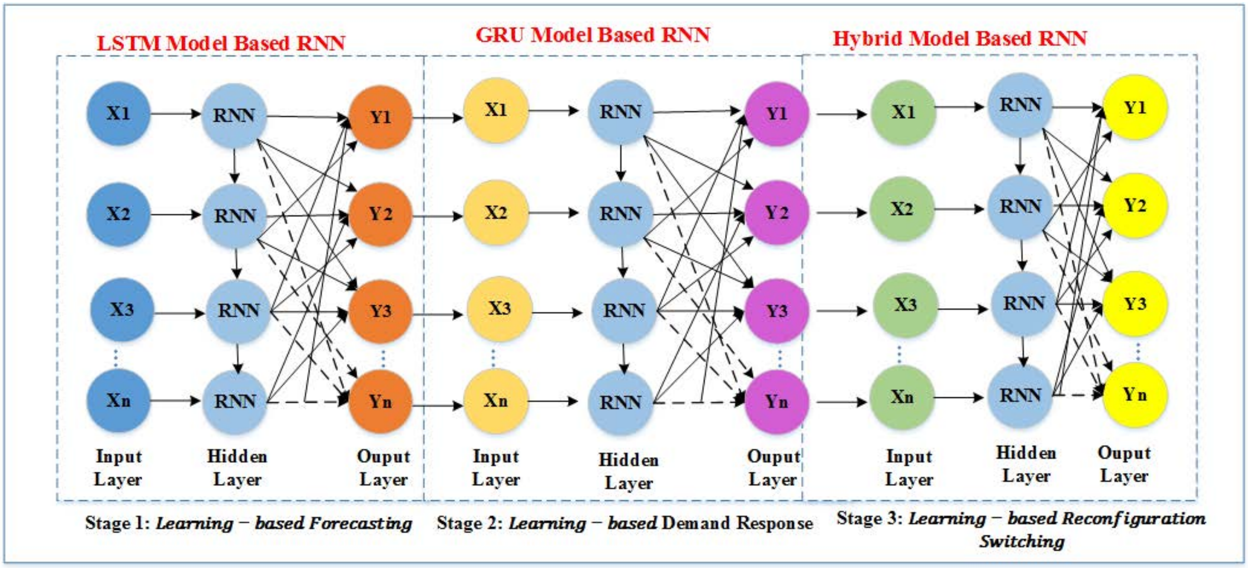

FIGURE 4. Proposed multi-stage learning-to-learning.

TABLE 8. Advantages and disadvantages of deep learning techniques on distribution automation system.

Technique	Advantages	Disadvantages
ANN	• Easy performance • Quick feed-forward manner • Complexity of low training time	• Restricted to supervised applications • Absence of the property coherence • Lack of spatial and temporal feature extraction
CNN	• Having acceptable accuracy • Correlating with the direction of the object of study • Distributed execution • Easy training manner applying gradient descent • Sparse data description • Detailed spatial point extraction	• High computing cost • It needs a large number of datasets to be effective • Restricted to supervised applications • needing high training memory complexity • High training time complexity • Failure to execute temporal data
LSTM	• Deriving precise temporal attributes • Extensible input data dimensions • Making shorter the pre-processing of information • Application for time series data	• Difficult in training process • Lack of spatial data modeling • High probability of overfitting • High sensibility to the initial case • Restricted to supervised usage • It can't accumulate to very deep models.
SAE	• Simple execution • Unsupervised feature extraction • Rapid feed-forward manner • Using the filters to apply dataset superior	• Large evaluation preference • Lack of spatial and temporal feature extraction • Absence of the property coherence • High probability of overfitting • needing extra training time
DBN	• Ability to model uncertainty • Unsupervised feature extraction • Small sample complexity	• Large training time complexity • Strong prior knowledge (conditional independence) • Lack of parameter convergence guarantee
GAN	• Ability to model uncertainty • Unsupervised feature extraction • Data Synthesis	• Large sample sophistication • Lack of parameter convergence verification • Limited variety • Decreased gradient • Untrustworthy estimations
AE	• Modeling uncertainties • Unsupervised feature extraction • Data Synthesis • Giving probabilistic classification and regression • Reliable estimation of the real probability distribution	• Large sample sophistication • Low sensitivity of evaluated distribution • Enormous testing time intricacy
DQN	• Resilience of the learning algorithm • Robustness for an uncertain environment	• Lack of convergence warranty • Slow convergence • Guessing deterministic policies

studies consulted in DL techniques have been compared by RMSE index in the forecasting field and by accuracy index in the security field. These criteria indexes directly refer to the performance of techniques. The accuracy index is one of the evaluation criteria that is mostly used in comparing studied methods in the security area. From Table. 9, it can be claimed

TABLE 9. Outcome/accuracy values in some studied applications to the comparison of different methods.

Ref	Technique	DL methods	Accuracy/Outcome	No
21	Load Forecasting	DBN-RBM	MAPE 0.21%	1
19		CNN, LSTM-RNN	MAPE 1.349%	2
24		DBN –RBS	MAPE 3%	3
28		CNN	RMSE 0.732%	4
29		LSTM-RNN	MAPE 8.18%	5
20		D-RNN	RMSE 111.9 kWh	6
46	Energy Forecasting	LSTM	RMSE 0.0642%	7
45		LSTM-RNN	RMSE 82.15%	8
39		CNN, LSTM	RMSE 0.37%,MAE 0.28%	9
37		DBN	MAPE 0.3%	10
44		RSDAE	RMSE 0.521%, MAE 0.213%	11
40		CNN-LSTM	RMSE 0.079%, MAE 0.056%	12
42		SAE-DBM	RMSE 0.094%, MAE 0.065%	13
49		DBN	MAPE 5.11%	14
50		DCNN	MAE 0.53%	15
55		CNN	MAPE 1.43%	16
58		DNN-TL	MAE 0.1%, RMSE 0.115%	17
59		LSTM-CNN	RMSE 17.90%	18
61		LSTM-DNN,GRU-DNN	MAPE 12.34%	19
66		LSTM	MAPE 12.74%	20
74	Security	SAE	Accuracy 95.78%	21
75		CNN-LSTM	Accuracy 89%	22
76		Deep CNN	Accuracy 99.3%	23
78		CNN	Accuracy 72%	24
83		CNN	Accuracy 97.8%	25
82		CNN	Accuracy 99.38%	26
84		2-D CNN	Accuracy 89%	27

with confidence that DL techniques along with the highest value of accuracy hold the most promising performance compared with other techniques. According to studies reviewed in recent years, it is obvious that CNNs have been utilized by most papers for detection. Based on the table. 6 the accuracy index is the most used and prevalent performance element by authors in the detection field. The network reconfiguration and restoration, voltage control, and demand response are the decision-making problems that are solved with DL-based Q-learning techniques. Therefore, Q-learning is the most famous and widely used technique in the decision-making area, as extracted from our investigations.

V. PROPOSED L2L TECHNIQUE FOR DAS

It is worth noting that hybrid methods had a remarkably high performance compared with single methods as they benefit from the advantages of several methods for providing the evaluation index. Also, studies have proven that RNN approaches such as LSTM have the best performance in the applications studied, such as the prediction domain. From studies, we devise a new intelligent architecture for real-time energy management of distribution networks using the multi-stage L2L approach. The optimal power flow (OPF) problem is applied to find the optimal scheduling of adjustable loads. However, altering the predicted values of load and renewable energies output power at any period can potentially impact the optimal scheduling of adjustable loads. Also, in emergency conditions, a fast response (real-time response) is essential to preserve the system from any possible load shedding. In the proposed architecture, instead of the conventional techniques that need to run the OPF problem, we used the real-time multi-stage L2L approach for scheduling the adjustable loads (such as PEV) as the load management. To this end, in the first stage, we used the DL technique, more specifically, the deep LSTM technique to forecast the real-time load and renewable energies output power values. In the second stage, we used the output power of the first stage as an input for the second stage, which is the optimal scheduling of adjustable loads at any time interval. In this stage, the GRU technique has been applied. To prevent the overloading and power outage by varying the topology of the network through some prelocated switches the third stage as the

TABLE 10. MSE for fully LSTM, GRU, and Hybrid LSTM-GRU for 3000 Epoch.

Learning Technique	Forecasted Value	MSE (kW)
LSTM	Market Price	0.0025
GRU		0.0023
Hybrid LSTM-GRU		0.0022
LSTM	Load Factor	0.0016
GRU		0.0022
Hybrid LSTM-GRU		0.0016
LSTM	Wind Turbine	0.0036
GRU		0.0027
Hybrid LSTM-GRU		0.0025
LSTM	PV	0.0032
GRU		0.0030
Hybrid LSTM-GRU		0.0030
LSTM	Smart Learning-based Network	0.0023
GRU		0.0020
Hybrid LSTM-GRU		0.0020

TABLE 11. Sensitivity Analysis of three different techniques (GRU, LSTM, and hybrid LSTM-GRU) considering 3000 epochs and Four layers of L2L.

No. Case	Hidden Layer Numbers		Output Layer	Market Price		Load Factor		Smart Learning-Based Network		PV		WT	
	No.1	No.2		MSE	Training Time (Min)	MSE	Training Time (Min)	MSE	Training Time (Min)	MSE	Training Time (Min)	MSE	Training Time (Min)
1	LSTM 100	LSTM 100	3-densely connected	0.0055	1.55	0.0053	1.53	0.0053	1.51	0.0054	1.54	0.0055	1.55
2	GRU 100	GRU 100	3-densely connected	0.006	1.51	0.0062	1.49	0.0059	1.48	0.0062	1.52	0.0059	1.52
3	GRU 100	LSTM 100	3-densely connected	0.0059	1.52	0.0061	1.51	0.0059	1.50	0.0061	1.53	0.0057	1.51
4	LSTM 200	LSTM 200	3-densely connected	0.0027	5.00	0.0025	4.55	0.0026	5.00	0.0025	4.47	0.0024	5.02
5	GRU 200	GRU 200	3-densely connected	0.0032	4.05	0.0031	4.14	0.0030	4.07	0.0028	3.58	0.0027	4.01
6	GRU 200	LSTM 200	3-densely connected	0.0028	4.25	0.0025	4.32	0.0027	4.16	0.0025	4.11	0.0026	4.19

reconfiguration switching stage is developed. Generally, the proposed model consists of several steps that connect different DL algorithms to solve modern power system problems, which has not been applied in any research until now. This model can simultaneously support different applications of DAS. The proposed multi-stage L2L framework follows three stages. (i) Forecasting stage: in the first stage, long short-term memory (LSTM) is used to forecast the hourly load demand, market price, and renewable energy power outputs. (ii) Demand response stage: in this stage, the gated recurrent unit (GRU) technique is utilized, where the input of this stage is the output of the previous stage. Also, the outputs of this stage are the optimal scheduling of adjustable loads, including both shiftable and curtailable loads, and (iii) smart learning-based reconfigurable switching stage: in this stage, the hybrid LSTM-GRU technique is employed, where the input of this stage is the load demand and generation units status which are the output of the first and second stage, while the output is the optimal reconfiguration switching status of the distribution grids. The structure of the proposed model is shown in Fig. 4. By utilizing the proposed technique, we cannot only manage the applications of the DAS such as the load and market operation of the network, but also we can increase the efficiency and reliability of the distribution grids. For the first time, we have developed a multi-stage deep learning based framework for forecasting, optimal adjustable loads scheduling, and optimal reconfiguration switching of distribution grids. To demonstrate the effectiveness and worthiness of the proposed framework, it has been tested on IEEE 33 bus test systems with modified reset capability. The network comprises three diesel generators (DGs), two wind turbines (WTs), and one photovoltaic (PV) system. The features of generators are summarized in Tables I2. Table 13 presents hourly normalized predicted values of WTs and PV generation, load, and market price.

After predicting the values of the hourly load factor, renewable energy output power (such as PV and WT), market price and smart learning-based network by LSTM, GRU, and Hybrid LSTM-GRU techniques and also optimal scheduling using a fully multi-stage learning technique, the mean square

TABLE 12. Features of generators.

Type	Min\Max Capacity (kW)	Cost($/kWh)	Min\Max of Power generation rate
DG1	500-1000	1.94	1500
DG2	400-1200	2.25	1000
DG3	100-2500	2.43	1500

errors (MSE) of three techniques per 3000 epochs have listed in Table 10. As can be seen in the table, the presented technique has the lowest MSE (highest accuracy) in three applications due to the highest number of epochs (3000).

As mentioned, the main goals of this survey are not only to provide deep learning applications in distribution automation systems but also to open a new approach/technique, called the collaborative learning-to-learning algorithm for multi-objective and complex distribution automation problems. These are a step above proposing hybrid techniques. Since the comparison of many techniques is impossible for a survey, we made some comparisons for different existing layers based on several combinations of LSTM and GRU, as shown in Table 11.

The number of hidden layers, the type of hidden layers, and more information about critical parameters of the deep learning technique is added to the paper in Table 11 for all three deep learning techniques. Moreover, for different layers, different parameter has been used as the error. More specifically, for market and load forecasting layers, prices are the main parameters to calculate the errors, while for renewable energies, PV and WT output power are the main parameters. Also, for the smart learning-based network, the voltages of buses are used for determining the error. It should be noted that all these cases have been run on the python program and also two 64-bit parallel software systems are considered to implement the proposed hybrid algorithm.

So in general, in the novel proposed model, at each step, LSTM and GRU algorithms, and their combination are used, respectively. These algorithms are a special structure of deep recursive neural networks that due to memory cells and control gates, have the ability to control the flow of information and determine the optimal time to remember and forget. GRU and LSTM algorithms, which are used in this model.

TABLE 13. Hourly normalized predicted values.

Hour	1-8							
Load (pu)	0.80	0.804	0.811	0.815	0.830	0.92	0.94	0.97
WTpower (pu)	0.115	0.115	0.115	0.115	0.115	0.022	0.114	0.092
PV power (pu)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Market($)	0.12	0.14	0.23	0.12	0.14	0.16	0.12	0.19
Hour	**9-16**							
Load (pu)	1	0.95	1	0.97	0.94	0.96	0.97	0.98
WTpower (pu)	0.119	0.203	0.209	0.302	0.390	0.385	0.214	0.152
PV power (pu)	0.0	0.005	0.019	0.125	0.152	0.201	0.201	0.210
Market($)	0.12	0.16	0.23	0.17	0.13	0.19	0.22	0.19
Hour	**17-24**							
Load (pu)	0.98	0.90	0.95	0.97	0.90	0.98	0.94	0.97
WTpower (pu)	0.119	0.115	0.100	0.095	0.184	0.119	0.021	0.021
PV power (pu)	0.180	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Market($)	0.22	0.21	0.19	0.17	0.19	0.12	0.13	0.16

They are more effective and more popular among researchers, especially for time series analysis and also has high accuracy and efficiency. In the first stage (forecasting stage), LSTM is used to predict hourly load demand, market price and renewable energy outputs, given that LSTM is very effective in forecasting time series. In the second stage (demand response stage), the GRU technique is used to achieve the optimal scheduling of adjustable loads, including removable and limited loads. Given that GRU has fewer parameters so it is trained faster, and it is also very efficient and effective for time series analysis. Also, according to studies obtained from the research, hybrid methods have a remarkably high performance compared with single methods. Therefore, in the third stage, the hybrid LSTM-GRU technique is employed. The main reason for selecting these two techniques is their performance, which is better than other deep learning approaches in time series prediction and high accuracy. As shown in Table 11, to compare the effectiveness of the proposed approach, we chose the GRU models and LSTM models as well as hybrid LSTM-GRU as we wanted to see if our proposed model enhances the overall performance and can outperform any of these algorithms. As shown, the GRU-LSTM model performs better because it has the optimal MSE index and training time compared to other hybrid algorithms.

VI. CONCLUSION

As the DAS transitions to a modern and advanced system, the conventional methods face limitations in analyzing, processing data and making decisions. Thus, DL methods are being developed and widely utilized in many applications fields of DAS with desirable results. From the methods viewpoint, they can be used in the forecast, detection, decision-making, etc. This paper presents a review of methods of DL in DAS applications (that is, load forecasting, energy forecasting, demand response, voltage control, energy market, reconfiguration and restoration, faults detection and cyber security). Also, the evaluation indexes have been investigated. A summary of literature reviews on DOS applications is offered in Table 7. Our future research will focus on a new perspective of DL called L2L algorithm that has great potential for DAS applications. The structure of the L2L platform that encompasses multi-stage is briefly described in this paper. A proposed L2L algorithm is a promising approach due to the fact that simultaneously supports different applications of DAS.

ACKNOWLEDGMENT

The financial support provided through the research project is highly acknowledged.

REFERENCES

[1] D. Gruenemeyer, "Distribution automation: How should it be evaluated?" in *Proc. Rural Electr. Power Conf.*, Apr. 1991, pp. 1–10.

[2] P. Parikh, "Distribution system automation," Dept. Elect. Comput. Eng., Univ. Western Ontario, Toronto, ON, Canada, ES 586b Course Project Rep., 2009.

[3] J. B. Bunch, "Guidelines for evaluating distribution automation," EPRI, Palo Alto, CA, USA, Tech. Rep. EL-3728, Nov. 1984.

[4] M. Z. Alom, T. M. Taha, C. Yakopcic, S. Westberg, P. Sidike, M. S. Nasrin, M. Hasan, B. C. Van Essen, A. A. S. Awwal, and V. K. Asari, "A state-of-the-art survey on deep learning theory and architectures," *Electron*, vol. 8, no. 3, p. 292, Mar. 2019.

[5] T. Bouwmans, S. Javed, M. Sultana, and S. K. Jung, "Deep neural network concepts for background subtraction: A systematic review and comparative evaluation," *J. Neural Netw.*, vol. 117, pp. 8–66, Sep. 2019.

[6] G. Yao, T. Lei, and J. Zhong, "A review of convolutional-neural-network-based action recognition," *Pattern Recognit. Lett.*, vol. 118, pp. 14–22, Feb. 2019.

[7] D. Güera and E. J. Delp, "Deepfake video detection using recurrent neural networks," in *Proc. 15th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Nov. 2018, pp. 1–6.

[8] E. López, C. Valle, H. Allende, and E. Gil, "Efficient training over long short-term memory networks for wind speed forecasting," in *Proc. Iberoamerican Congr. Pattern Recognit.* Cham, Switzerland: Springer, Nov. 2016, pp. 409–416.

[9] M. A. Wani, F. A. Bhat, S. Afzal, and A. I. Khan, "Introduction to deep learning," in *Advances in Deep Learning*, vol. 57. Singapore: Springer, Mar. 2020, pp. 1–11.

[10] A. K. Ozcanli, F. Yaprakdal, and M. Baysal, "Deep learning methods and applications for electrical power systems: A comprehensive review," *Int. J. Energy Res.*, vol. 44, no. 9, pp. 7136–7157, Jul. 2020.

[11] M. Jafari, A. Kavousi-Fard, T. Niknam, and O. Avatefipour, "Stochastic synergies of urban transportation system and smart grid in smart cities considering V2G and V2S concepts," *Energy*, vol. 215, Jan. 2021, Art. no. 119054.

[12] S. Khatoon and A. K. Singh, "Effects of various factors on electric load forecasting: An overview," in *Proc. 6th IEEE Power India Int. Conf. (PIICON)*, Dec. 2014, pp. 1–5.

[13] A. Almalaq and G. Edwards, "A review of deep learning methods applied on load forecasting," in *Proc. 16th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Cancun, Mexico, Dec. 2017, pp. 511–516.

[14] M. A. M. Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah, and F. Hussin, "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review," *Renew. Sustain. Energy Rev.*, vol. 70, pp. 1108–1118, Apr. 2017.

[15] Z. Guo, K. Zhou, X. Zhang, and S. Yang, "A deep learning model for short-term power load and probability density forecasting," *Energy*, vol. 160, pp. 1186–1200, Oct. 2018.

[16] G. M. U. Din and A. K. Marnerides, "Short term power load forecasting using deep neural networks," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Jan. 2017, pp. 594–598.
[17] X. Dong, L. Qian, and L. Huang, "Short-term load forecasting in smart grid: A combined CNN and K-means clustering approach," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Feb. 2017, pp. 119–125.
[18] W. He, "Load forecasting via deep neural networks," *Proc. Comput. Sci.*, vol. 122, pp. 308–314, Dec. 2017.
[19] Y. He, J. Deng, and H. Li, "Short-term power load forecasting with deep belief network and copula models," in *Proc. 9th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. (IHMSC)*, vol. 1, Aug. 2017, pp. 191–194.
[20] A. Dedinec, S. Filiposka, A. Dedinec, and L. Kocarev, "Deep belief network based electricity load forecasting: An analysis of Macedonian case," *Energy*, vol. 115, pp. 1688–1700, Nov. 2016.
[21] L. Wen, K. Zhou, and S. Yang, "Load demand forecasting of residential buildings using a deep learning model," *Electr. Power Syst. Res.*, vol. 179, Feb. 2020, Art. no. 106073.
[22] B. Jatin and D. Toshniwal, "Empirical mode decomposition based deep learning for electricity demand forecasting," *IEEE Access*, vol. 6, pp. 49144–49156, 2018.
[23] X. Qiu, Y. Ren, P. N. Suganthan, and G. A. J. Amaratunga, "Empirical mode decomposition based ensemble deep learning for load demand time series forecasting," *Appl. Soft Comput.*, vol. 54, pp. 246–255, May 2017.
[24] A. Estebsari and R. Rajabi, "Single residential load forecasting using deep learning and image encoding techniques," *Electronics*, vol. 9, no. 1, p. 68, Jan. 2020.
[25] C. Tong, J. Li, C. Lang, F. Kong, J. Niu, and J. J. P. C. Rodrigues, "An efficient deep model for day-ahead electricity load forecasting with stacked denoising auto-encoders," *J. Parallel Distrib. Comput.*, vol. 117, pp. 267–273, Jul. 2018.
[26] D. L. Marino, K. Amarasinghe, and M. Manic, "Building energy load forecasting using deep neural networks," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2016, pp. 7046–7051.
[27] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," in *Proc. IEEE 26th Int. Symp. Ind. Electron. (ISIE)*, Jun. 2017, pp. 1483–1488.
[28] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu, and Y. Zhang, "Short-term residential load forecasting based on LSTM recurrent neural network," *IEEE Trans. Smart Grid*, vol. 10, no. 1, pp. 841–851, Sep. 2017.
[29] H. Shi, M. Xu, and R. Li, "Deep learning for household load forecasting—A novel pooling deep RNN," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 5271–5280, Sep. 2018.
[30] M. Tan, S. Yuan, S. Li, Y. Su, and H. Li, and F. He, "Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 2937–2948, Jul. 2020.
[31] J. Moon, S. Jung, J. Rew, S. Rho, and E. Hwang, "Combination of short-term load forecasting models based on a stacking ensemble approach," *Energy Buildings*, vol. 216, Jun. 2020, Art. no. 109921.
[32] Y. Chen, Y. Wang, D. S. Kirschen, and B. Zhang, "Model-free renewable scenario generation using generative adversarial networks," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3265–3275, May 2018.
[33] M. Khodayar, J. Wang, and M. Manthouri, "Interval deep generative neural network for wind speed forecasting," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 3974–3989, Jul. 2019.
[34] M. Khodayar, S. Mohammadi, M. E. Khodayar, J. Wang, and G. Liu, "Convolutional graph autoencoder: A generative deep neural network for probabilistic spatio-temporal solar irradiance forecasting," *IEEE Trans. Sustain. Energy*, vol. 11, no. 2, pp. 571–583, Apr. 2020.
[35] H. Z. Wang, G. B. Wang, G. Q. Li, J. C. Peng, and Y. T. Liu, "Deep belief network based deterministic and probabilistic wind speed forecasting approach," *Appl. Energy*, vol. 182, pp. 80–93, Nov. 2016.
[36] H. Z. Wang, G.-Q. Li, G.-B. Wang, J.-C. Peng, H. Jiang, and Y.-T. Liu, "Deep learning based ensemble approach for probabilistic wind power forecasting," *Appl. Energy*, vol. 188, pp. 56–70, Feb. 2017.
[37] H. Liu, X.-W. Mi, and Y.-F. Li, "Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network," *Energy Convers. Manage.*, vol. 156, pp. 498–514, Jan. 2018.
[38] W. Wu, K. Chen, Y. Qiao, and Z. Lu, "Probabilistic short-term wind power forecasting based on deep neural networks," in *Proc. Int. Conf. Probabilistic Methods Appl. Power Syst. (PMAPS)*, Oct. 2016, pp. 1–8.
[39] K. Wang, X. Qi, H. Liu, and J. Song, "Deep belief network based K-means cluster approach for short-term wind power forecasting," *Energy*, vol. 165, pp. 840–852, Dec. 2018.
[40] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, 2020.
[41] A. T. Sergio and T. B. Ludermir, "Deep learning for wind speed forecasting in northeastern region of Brazil," in *Proc. Brazilian Conf. Intell. Syst. (BRACIS)*, Nov. 2015, pp. 322–327.
[42] M. Khodayar, O. Kaynak, and M. E. Khodayar, "Rough deep neural architecture for short-term wind speed forecasting," *IEEE Trans. Ind. Informat.*, vol. 13, no. 6, pp. 2770–2779, Dec. 2017.
[43] M. Abdel-Nasser and K. Mahmoud, "Accurate photovoltaic power forecasting models using deep LSTM-RNN," *Neural Comput. Appl.*, vol. 31, no. 7, pp. 2727–2740, Oct. 2017.
[44] A. Gensler, J. Henze, B. Sick, and N. Raabe, "Deep learning for solar power forecasting—An approach using autoencoder and LSTM neural networks," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 002858–002865.
[45] L. Wen, K. Zhou, S. Yang, and X. Lu, "Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting," *Energy*, vol. 171, pp. 1053–1065, Mar. 2019.
[46] C.-J. Huang and P.-H. Kuo, "Multiple-input deep convolutional neural network model for short-term photovoltaic power forecasting," *IEEE Access*, vol. 7, pp. 74822–74834, 2019.
[47] L. L. Li, P. Cheng, H. C. Lin, and H. Dong, "Short-term output power forecasting of photovoltaic systems based on the deep belief net," *Adv. Mech. Eng.*, vol. 9, no. 9, Sep. 2017, Art. no. 1687814017715983.
[48] H. Wang, H. Yi, J. Peng, G. Wang, Y. Liu, and H. Jiang, "Deterministic and probabilistic forecasting of photovoltaic power based on deep convolutional neural network," *Energy Convers. Manage.*, vol. 153, pp. 409–422, Dec. 2017.
[49] H. Zang, L. Cheng, T. Ding, K. W. Cheung, Z. Liang, Z. Wei, and G. Sun, "Hybrid method for short-term photovoltaic power forecasting based on deep convolutional neural network," *IET Gener., Transm. Distrib.*, vol. 12, no. 20, pp. 4557–4567, Nov. 2018.
[50] Z. Zhen, J. Liu, Z. Zhang, F. Wang, H. Chai, Y. Yu, X. Lu, T. Wang, and Y. Lin, "Deep learning based surface irradiance mapping model for solar PV power forecasting using sky image," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3385–3396, Aug. 2020.
[51] S. M. J. Jalali, S. Ahmadian, A. Kavousi-Fard, A. Khosravi, and S. Nahavandi, "Automated deep CNN-LSTM architecture design for solar irradiance forecasting," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 1, pp. 54–65, Jan. 2022.
[52] H. Aprillia, H. T. Yang, and C. M. Huang, "Short-term photovoltaic power forecasting using a convolutional neural network–salp swarm algorithm," *Energies*, vol. 13, no. 8, p. 1879, 2020.
[53] P. Li, K. Zhou, X. Lu, and S. Yang, "A hybrid deep learning model for short-term PV power forecasting," *Appl. Energy*, vol. 259, Feb. 2020, Art. no. 114216.
[54] M. Mohammadi, S. Z. Tajalli, and K. Esapour, "Effective demand response and GANs for optimal constraint unit commitment in solar-tidal based microgrids," *IET Renew. Power Gener.*, pp. 1–11, Nov. 2021, doi: 10.1049/rpg2.12331.
[55] A. S. Qureshi, A. Khan, A. Zameer, and A. Usman, "Wind power prediction using deep neural network based meta regression and transfer learning," *Appl. Soft Comput.*, vol. 58, pp. 742–755, Sep. 2017.
[56] P.-H. Kuo and C.-J. Huang, "An electricity price forecasting model by hybrid structured deep neural networks," *Sustainability*, vol. 10, no. 4, p. 1280, Apr. 2018.
[57] U. Ugurlu, I. Oksuz, and O. Tas, "Electricity price forecasting using recurrent neural networks," *Energies*, vol. 11, no. 5, p. 1255, May 2018.
[58] J. Lago, F. De Ridder, and B. De Schutter, "Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms," *Appl. Energy*, vol. 221, pp. 386–405, Jul. 2018.
[59] L. Wang, Z. Zhang, and J. Chen, "Short-term electricity price forecasting with stacked denoising autoencoders," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2673–2681, Jul. 2017.
[60] S. Mujeeb, N. Javaid, M. Akbar, R. Khalid, O. Nazeer, and M. Khan, "Big data analytics for price and load forecasting in smart grids," in *Proc. Int. Conf. Broadband Wireless Comput., Commun. Appl.*, Oct. 2018, pp. 77–87.

[61] C. Zhang, R. Li, H. Shi, and F. Li, "Deep learning for day-ahead electricity price forecasting," *IET Smart Grid*, vol. 3, no. 4, pp. 462–469, Aug. 2020.
[62] A. Marszałek and T. Burczynski, "Forecasting day-ahead spot electricity prices using deep neural networks with attention mechanism," *J. Smart Environments Green Comput.*, vol. 1, no. 1, pp. 21–31, 2021.
[63] A. A. Elsaeidy, N. Jagannath, A. G. Sanchis, A. Jamalipour, and K. S. Munasinghe, "Replay attack detection in smart cities using deep learning," *IEEE Access*, vol. 8, pp. 137825–137837, 2020.
[64] A. A. Diro and N. Chilamkurti, "Distributed attack detection scheme using deep learning approach for Internet of Things," *Future Gener. Comput. Syst.*, vol. 82, pp. 761–768, May 2018.
[65] Y. Otoum, D. Liu, and A. Nayak, "DL-IDS: A deep learning–based intrusion detection framework for securing IoT," *Trans. Emerg. Telecommun. Technol.*, vol. 30, no. 11, p. 3803, Nov. 2019.
[66] G. Thamilarasu and S. Chawla, "Towards deep-learning-driven intrusion detection for the Internet of Things," *Sensors*, vol. 19, no. 9, p. 1977, Apr. 2019.
[67] B. A. Tama and K.-H. Rhee, "Attack classification analysis of IoT network via deep learning approach," *Res. Briefs Inf. Commun. Technol. Evol.*, vol. 3, pp. 1–9, Nov. 2017.
[68] M. Aloqaily, S. Otoum, I. A. Ridhawi, and Y. Jararweh, "An intrusion detection system for connected vehicles in smart cities," *Ad Hoc Netw.*, vol. 90, Jul. 2019, Art. no. 101842.
[69] H. Wang, J. Ruan, G. Wang, B. Zhou, Y. Liu, X. Fu, and J. Peng, "Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks," *IEEE Trans. Ind. Informat.*, vol. 14, no. 11, pp. 4766–4778, Nov. 2018.
[70] T.-E. Huang, Q. Guo, H. Sun, C.-W. Tan, and T. Hu, "A deep spatial-temporal data-driven approach considering microclimates for power system security assessment," *Appl. Energy*, vol. 237, pp. 36–48, Mar. 2019.
[71] M. N. Hasan, R. N. Toma, A.-A. Nahid, M. M. M. Islam, and J.-M. Kim, "Electricity theft detection in smart grid systems: A CNN-LSTM based approach," *Energies*, vol. 12, no. 17, p. 3310, 2019.
[72] R. R. Bhat, R. D. Trevizan, R. Sengupta, X. Li, and A. Bretas, "Identifying nontechnical power loss via spatial and temporal deep learning," in *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, vol. 18, Dec. 2016, pp. 272–279.
[73] Y. 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 Trans. Smart Grid*, vol. 8, no. 5, pp. 2505–2516, Sep. 2017.
[74] Z. Zheng, Y. Yang, X. Niu, H. N. Dai, and Y. Zhou, "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1606–1615, Apr. 2018.
[75] H. M. Rouzbahani, H. Karimipour, and L. Lei, "An ensemble deep convolutional neural network model for electricity theft detection in smart grids," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2020, pp. 3637–3642.
[76] A. Sayghe, J. Zhao, and C. Konstantinou, "Evasion attacks with adversarial deep learning against power system state estimation," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2020, pp. 1–5.
[77] Y. Zhang, J. Wang, and B. Chen, "Detecting false data injection attacks in smart grids: A semi-supervised deep learning approach," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 623–634, Jan. 2021.
[78] H. Yang, L. Cheng, and M. C. Chuah, "Deep-learning-based network intrusion detection for SCADA systems," in *Proc. IEEE Conf. Commun. Netw. Secur. (CNS)*, Jun. 2019, pp. 1–7.
[79] J. M. H. Arteaga, F. Hancharou, and F. Thams, and S. Chatzivasileiadis, "Deep learning for power system security assessment," in *Proc. IEEE Milan PowerTech*, Jun. 2019, pp. 1–6.
[80] C. Xu, K. Zhang, and J. Li, "Detect the electricity theft event using text CNN," in *Proc. IOP Conf., Earth Environ. Sci.*, 2021, vol. 675, no. 1, Art. no. 012136.
[81] Z. Aslam, N. Javaid, A. Ahmad, A. Ahmed, and S. M. Gulfam, "A combined deep learning and ensemble learning methodology to avoid electricity theft in smart grids," *Energies*, vol. 13, no. 21, p. 5599, 2020.
[82] S. Ahmadian, H. Malki, and Z. Han, "Cyber attacks on smart energy grids using generative adverserial networks," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2018, pp. 942–946.
[83] Y. Wang, M. Liu, and Z. Bao, "Deep learning neural network for power system fault diagnosis," in *Proc. 35th Chin. Control Conf.*, Jul. 2016, pp. 6678–6683.
[84] S. Zhang, Y. Wang, M. Liu, and Z. Bao, "Data-based line trip fault prediction in power systems using LSTM networks and SVM," *IEEE Access*, vol. 6, pp. 7675–7686, 2018.
[85] Y. Wang, M. Liu, Z. Bao, and S. Zhang, "Stacked sparse autoencoder with PCA and SVM for data-based line trip fault diagnosis in power systems," *Neural Comput. Appl.*, vol. 31, no. 10, pp. 6719–6731, Apr. 2019.
[86] D. Mnyanghwalo, H. Kundaeli, E. Kalinga, and N. Hamisi, "Deep learning approaches for fault detection and classifications in the electrical secondary distribution network: Methods comparison and recurrent neural network accuracy comparison," *Cogent Eng.*, vol. 7, no. 1, Jan. 2020, Art. no. 1857500.
[87] S. Chan, I. Oktavianti, V. Puspita, and P. Nopphawan, "Convolutional adversarial neural network (CANN) for fault diagnosis within a power system : Addressing the challenge of event correlation for diagnosis by power disturbance monitoring equipment in a smart grid," in *Proc. Int. Conf. Inf. Commun. Technol. (ICOIACT)*, Jul. 2019, pp. 596–601.
[88] F. Rudin, G. J. Li, and K. Wang, "An algorithm for power system fault analysis based on convolutional deep learning neural network," *Int. J. Res. Educ. Sci. Methods (IJARESM)*, vol. 5, no. 9, pp. 11–18, 2017.
[89] L. Wang, X. Zhao, J. Pei, and G. Tang, "Transformer fault diagnosis using continuous sparse autoencoder," *SpringerPlus*, vol. 5, no. 1, pp. 1–13, Dec. 2016.
[90] Z. Xu, W. Mo, Y. Wang, S. Luo, and T. Liu, "Transformer fault diagnosis based on deep brief sparse autoencoder," in *Proc. Chin. Control Conf. (CCC)*, Jul. 2019, pp. 7432–7435.
[91] J. Dai, H. Song, G. Sheng, and X. Jiang, "Dissolved gas analysis of insulating oil for power transformer fault diagnosis with deep belief network," *IEEE Trans. Dielectr. Elect. Insul.*, vol. 24, no. 5, pp. 2828–2835, Oct. 2017.
[92] J. Q. James, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1694–1703, Mar. 2019.
[93] X. Qin, Y. Zhang, W. Mei, G. Dong, J. Gao, P. Wang, J. Deng, and H. Pan, "A cable fault recognition method based on a deep belief network," *Comput. Elect. Eng.*, vol. 71, vol. 71, pp. 452–464, Aug. 2018.
[94] J. Liang, T. Jing, H. Niu, and J. Wang, "Two-terminal fault location method of distribution network based on adaptive convolution neural network," *IEEE Access*, vol. 8, pp. 54035–54043, 2020.
[95] M. F. Guo, X. D. Zeng, D. Y. Chen, and N. C. Yang, "Deep-learning-based earth fault detection using continuous wavelet transform and convolutional neural network in resonant grounding distribution systems," *IEEE Sensors J.*, vol. 18, no. 3, pp. 1291–1300, Feb. 2018.
[96] M. Jimenez Aparicio, "Fast fault location method for a distribution system with high penetration of PV," Ph.D. dissertation, Georgia Inst. Technol., Atlanta, GA, USA, 2020.
[97] M. F. Guo, N. C. Yang, and W. F. Chen, "Deep-learning-based fault classification using Hilbert–Huang transform and convolutional neural network in power distribution systems," *IEEE Sensors J.*, vol. 19, no. 16, pp. 6905–6913, Aug. 2019.
[98] J. Liu, F. Qu, X. Hong, and H. Zhang, "A small-sample wind turbine fault detection method with synthetic fault data using generative adversarial nets," *IEEE Trans. Ind. Informat.*, vol. 15, no. 7, pp. 3877–3888, Jul. 2019.
[99] C. Ren and Y. Xu, "Transfer learning-based power system online dynamic security assessment: Using one model to assess many unlearned faults," *IEEE Trans. Power Syst.*, vol. 35, no. 1, pp. 821–824, Jan. 2020.
[100] Q. F. Zhou, Y. Ge, X. D. Song, K. Lai, and J. C. Guo, "Fault diagnosis of distribution terminal units' measurement system based on generative adversarial network combined with convolutional neural network," in *Proc. IOP Conf., Mater. Sci. Eng.*, 2020, vol. 752, no. 1, Art. no. 012016.
[101] R. S. Rao, K. Ravindra, K. Satish, and S. V. L. Narasimham, "Power loss minimization in distribution system using network reconfiguration in the presence of distributed generation," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 317–325, Feb. 2013.
[102] T. Niknam, A. K. Fard, and A. Seifi, "Distribution feeder reconfiguration considering fuel cell/wind/photovoltaic power plants," *Renew. Energy*, vol. 37, no. 1, pp. 213–225, Jan. 2012.
[103] S. Zhang, H. Cheng, D. Wang, L. Zhang, F. Li, and L. Yao, "Distributed generation planning in active distribution network considering demand side management and network reconfiguration," *Appl. Energy*, vol. 228, pp. 1921–1936, Oct. 2018.
[104] S. F. Santos, D. Z. Fitiwi, M. R. M. Cruz, C. M. P. Cabrita, and J. P. S. Catalão, "Impacts of optimal energy storage deployment and network reconfiguration on renewable integration level in distribution systems," *Appl. Energy*, vol. 185, pp. 44–55, Jan. 2017.

[105] L. Bai, T. Jiang, F. Li, H. Chen, and X. Li, "Distributed energy storage planning in soft open point based active distribution networks incorporating network reconfiguration and DG reactive power capability," *Appl. Energy*, vol. 210, pp. 1082–1091, Jan. 2018.
[106] M. A. Kashem, V. Ganapathy, and G. B. Jasmon, "Network reconfiguration for load balancing in distribution networks," *IEE Proc.-Gener., Transmiss. Distrib.*, vol. 146, no. 6, pp. 563–567, Nov. 1999.
[107] A. Kavousi-Fard and M. R. Akbari-Zadeh, "Reliability enhancement using optimal distribution feeder reconfiguration," *Neurocomputing*, vol. 106, pp. 1–11, Apr. 2013.
[108] A. Kavousi-Fard and T. Niknam, "Multi-objective stochastic distribution feeder reconfiguration from the reliability point of view," *Energy*, vol. 64, pp. 342–354, Jan. 2014.
[109] Y. Gao, W. Wang, J. Shi, and N. Yu, "Batch-constrained reinforcement learning for dynamic distribution network reconfiguration," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 5357–5369, Nov. 2020.
[110] Y. Gao, J. Shi, W. Wang, and N. Yu, "Dynamic distribution network reconfiguration using reinforcement learning," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Beijing, China, Oct. 2019, pp. 1–7.
[111] X. Ji, Z. Yin, Y. Zhang, B. Xu, and Q. Liu, "Real-time autonomous dynamic reconfiguration based on deep learning algorithm for distribution network," *Electr. Power Syst. Res.*, vol. 195, Jun. 2021, Art. no. 107132.
[112] D. Ye, M. Zhang, and D. Sutanto, "A hybrid multiagent framework with Q-learning for power grid systems restoration," *IEEE Trans. Power Syst.*, vol. 26, no. 4, pp. 2434–2441, Nov. 2011.
[113] S. Das, S. Bose, S. Pal, N. N. Schulz, C. M. Scoglio, and B. Natarajan, "Dynamic reconfiguration of shipboard power systems using reinforcement learning," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 669–676, May 2013.
[114] J. Hong, "A multiagent Q-learning-based restoration algorithm for resilient distribution system operation," M.S. thesis, Univ. Central Florida, Orlando, FL, USA, 2017. [Online]. Available: https://stars.library.ucf.edu/etd/5572
[115] L. R. Ferreira, A. R. Aoki, and G. Lambert-Torres, "A reinforcement learning approach to solve service restoration and load management simultaneously for distribution networks," *IEEE Access*, vol. 7, pp. 145978–145987, 2019.
[116] J. G. Vlachogiannis and N. Hatziargyriou, "Reinforcement learning (RL) to optimal reconfiguration of radial distribution system (RDS)," in *Proc. Hellenic Conf. Artif. Intell.*, 2004, pp. 439–446.
[117] J. G. Vlachogiannis and N. D. Hatziargyriou, "Reinforcement learning for reactive power control," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1317–1325, Aug. 2004.
[118] Y. Xu, W. Zhang, W. Liu, and F. Ferrese, "Multiagent-based reinforcement learning for optimal reactive power dispatch," *IEEE Trans. Syst., Man, Cybern., C (Appl. Rev.)*, vol. 42, no. 6, pp. 1742–1751, Nov. 2012.
[119] H. Xu, A. D. Dominguez-Garcia, and P. W. Sauer, "Optimal tap setting of voltage regulation transformers using batch reinforcement learning," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 1990–2001, May 2020.
[120] W. Wang, N. Yu, Y. Gao, and J. Shi, "Safe off-policy deep reinforcement learning algorithm for volt-VAR control in power distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3008–3018, Jul. 2020.
[121] Q. Yang, G. Wang, A. Sadeghi, G. B. Giannakis, and J. Sun, "Two-timescale voltage control in distribution grids using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2313–2323, May 2020.
[122] Y. Zhang, X. Wang, J. Wang, and Y. Zhang, "Deep reinforcement learning based volt-VAR optimization in smart distribution systems," *IEEE Trans. Smart Grid*, vol. 12, no. 1, pp. 361–371, Jan. 2021.
[123] Y. Gao, W. Wang, and N. Yu, "Consensus multi-agent reinforcement learning for volt-VAR control in power distribution networks," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3594–3604, Jul. 2021.
[124] W. Wang, N. Yu, J. Shi, and Y. Gao, "Volt-VAR control in power distribution systems with deep reinforcement learning," in *Proc. IEEE Int. Conf. Commun., Control, Comput. Technol. Smart Grids (SmartGridComm)*, Oct. 2019, pp. 1–7.
[125] C. Li, C. Jin, and R. Sharma, "Coordination of PV smart inverters using deep reinforcement learning for grid voltage regulation," in *Proc. 18th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2019.
[126] L. Yin, C. Zhang, Y. Wang, F. Gao, J. Yu, and L. Cheng, "Emotional deep learning programming controller for automatic voltage control of power systems," *IEEE Access*, vol. 9, pp. 31880–31891, 2021.
[127] R. Lu and S. H. Hong, "Incentive-based demand response for smart grid with reinforcement learning and deep neural network," *Appl. Energy*, vol. 236, pp. 937–949, Feb. 2019.
[128] X. Y. Zhang, "A data-driven approach for coordinating air conditioning units in buildings during demand response events," Virginia Polytech. Inst., State Univ. Virginia, Blacksburg, VA, USA, 2019.
[129] J. Hao, "Multi-agent reinforcement learning embedded game for the optimization of building energy control and power system planning," 2019, *arXiv:1901.07333*.
[130] A. K. Sharma and A. Saxena, "A demand side management control strategy using whale optimization algorithm," *SN Appl. Sci.*, vol. 1, no. 8, pp. 1–15, Aug. 2019.
[131] A. Ghasemkhani and L. Yang, "Reinforcement learning based pricing for demand response," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–6.
[132] T. Remani, E. A. Jasmin, and T. P. I. Ahamed, "Residential load scheduling with renewable generation in the smart grid: A reinforcement learning approach," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3283–3294, Sep. 2019.
[133] A. Chiş, J. Lundén, and V. Koivunen, "Reinforcement learning-based plug-in electric vehicle charging with forecasted price," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 3674–3684, May 2017.
[134] R. Lu, S. H. Hong, and X. Zhang, "A dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach," *Appl. Energy*, vol. 220, pp. 220–230, Jun. 2018.
[135] S. Najafi, S. Talari, A. S. Gazafroudi, M. Shafie-Khah, J. M. Corchado, and J. P. S. Catalao, "Decentralized control of DR using a multi-agent method," in *Sustainable Interdependent Networks*, M. H. Amini, K. G. Boroojeni, S. S. Iyengar, P. M. Pardalos, F. Blaabjerg, and A. M. Madni, Eds. Cham, Switzerland: Springer, 2018, pp. 233–249.
[136] X. Huang, S. H. Hong, M. Yu, Y. Ding, and J. Jiang, "Demand response management for industrial facilities: A deep reinforcement learning approach," *IEEE Access*, vol. 7, pp. 82194–82205, 2019.
[137] J. R. Vázquez-Canteli and Z. Nagy, "Reinforcement learning for demand response: A review of algorithms and modeling techniques," *Appl. Energy*, vol. 235, pp. 1072–1089, Feb. 2019.

MINA JAFARI was born in Shiraz, Iran. She received the B.Sc. degree from Zabol University, Zabol, Iran, and the M.Sc. degree from the Shiraz University of Technology, Shiraz, Iran, where she is currently pursuing the Ph.D. degree in electrical power engineering. Her current research interests include stochastic analysis, evolutionary algorithms, operation, management, and cyber security analysis of smart grids, smart city, artificial intelligence, machine learning, and digital twin.

ABDOLLAH KAVOUSI-FARD (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from the Shiraz University of Technology, Shiraz, Iran, in 2009, the M.Sc. degree in electrical engineering from Shiraz University, Shiraz, in 2011, and the Ph.D. degree in electrical engineering from the Shiraz University of Technology, in 2016. He was a Researcher with the University of Denver, Denver, CO, USA, from 2015 to 2016, conducting research on microgrids. He was a Postdoctoral Research Assistant at the University of Michigan, MI, USA, from 2016 to 2018. He is currently an Assistant Professor at the Shiraz University of Technology. He has published more than 100 research papers in prestigious international journals and peer-reviewed conference proceedings. His current research interests include operation, management and cyber security analysis of smart grids, microgrid, smart city, electric vehicles, artificial intelligence, and machine learning.

MORTEZA DABBAGHJAMANESH (Senior Member, IEEE) received the M.Sc. degree in electrical engineering from Northern Illinois University, DeKalb, IL, USA, in 2014, and the Ph.D. degree in electrical and computer engineering from Louisiana State University, Baton Rouge, LA, USA, in 2019. In 2015, he joined the Renewable Energy and Smart Grid Laboratory, Louisiana State University. In 2019, he joined the Design and Optimization of Energy Systems (DOES) Laboratory, The University of Texas at Dallas, Richardson, TX, USA. He is currently a Research Associate at The University of Texas at Dallas. His current research interests include power system operation, reliability, and resiliency, renewable energy sources, cybersecurity analysis, machine learning, smart grids, and microgrids.

MAZAHER KARIMI (Senior Member, IEEE) received the Ph.D. degree in electrical energy and power system from the University of Malaya, Kuala Lumpur, Malaysia, in 2013.

He has worked as a Research Associate with The University of Manchester, from 2016 to 2017. From 2017 to 2020, he was an Assistant Professor at Gonbad Kavous University, Iran. He is currently an Assistant Professor with the School of Technology and Innovations, University of Vaasa, Vaasa, Finland. His current research interests include smart grid applications; wide-area monitoring, protection, and control; distributed generation; and power system stability.