Deep Learning-Based Fault Diagnosis of Photovoltaic Systems: A Comprehensive Review and Enhancement Prospects

MAJDI MANSOURI[1], (Senior Member, IEEE),
MOHAMED TRABELSI[2], (Senior Member, IEEE),
HAZEM NOUNOU[1], (Senior Member, IEEE),
AND MOHAMED NOUNOU[3], (Senior Member, IEEE)

[1]Electrical and Computer Engineering Program, Texas A&M University at Qatar, Doha, Qatar
[2]Electronic and Communications Engineering Department, Kuwait College of Science and Technology, Safat, Kuwait City 13133, Kuwait
[3]Chemical Engineering Program, Texas A&M University at Qatar, Doha, Qatar

Corresponding author: Majdi Mansouri (majdi.mansouri@qatar.tamu.edu)

ABSTRACT Photovoltaic (PV) systems are subject to failures during their operation due to the aging effects and external/environmental conditions. These faults may affect the different system components such as PV modules, connection lines, converters/inverters, which can lead to a decrease in the efficiency, performance, and further system collapse. Thus, a key factor to be taken into consideration in high-efficiency grid-connected PV systems is the fault detection and diagnosis (FDD). The performance of the FDD method depends mainly on the quality of the extracted features including real-time changes, phase changes, trend changes, and faulty modes. Thus, the data representation learning is the core stage of intelligent FDD techniques. Recently, due to the enhancement of computing capabilities, the increase of the big data use, and the development of effective algorithms, the deep learning (DL) tool has witnessed a great success in data science. Therefore, this paper proposes an extensive review on deep learning based FDD methods for PV systems. After a brief description of the DL-based strategies, techniques for diagnosing PV systems proposed in recent literature are overviewed and analyzed to point out their differences, advantages and limits. Future research directions towards the improvement of the performance of the DL-based FDD techniques are also discussed. This review paper aims to systematically present the development of DL-based FDD for PV systems and provide guidelines for future research in the field.

INDEX TERMS Fault diagnosis, deep learning, photovoltaic systems.

I. INTRODUCTION

Photovoltaic (PV)-based electrical power generation has been a growing research area in the academia and industry fields [1], [2], where the grid-connected PV systems have witnessed the highest growth rate. Therefore, the high-performance/reliability operation of PV systems has become a top priority. PV systems' faults can be divided, according to their time characteristics, into three major categories: intermittent, abrupt, or incipient faults. Temporary or intermittent faults refer to faults that clear or change over time such as

partial shading or environmental stress (e.g., dust or contamination). Permanent or abrupt faults refer to faults that occur instantaneously often as a result of a damage to the PV array such as line-to-line or line-to-ground short circuits, junction box faults, connector disconnection, open-circuit faults, and hot spots. Incipient faults are considered as the most challenging failures due to their small amplitudes and slow dynamics. If not detected at an early stage, they can result in gradual damage to the PV cells leading to serious faults. Incipient faults can occur in both DC and AC sides. Examples of DC-side (PV modules and DC/DC converter) incipient faults are PV module defects such as yellowing and browning of the solar cells, delamination, bubbles, cracks, gaps, and defects

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

in the anti-reflective coating. The AC-side (inverter and grid side) faults include Insulated Gate Bipolar Transistor (IGBT) faults, wiring degradation, aging, islanding, and overheating.

Therefore, it is essential to develop enhanced FDD algorithms aiming at increasing the reliability and efficiency of PV systems [3]. With the rapid development of information and automation technologies, the demand and requirements for FDD algorithms are increasing, and data-driven process control methods are also being continuously developed and improved. Thanks to the powerful representation learning ability of DL algorithms, intelligent FDD become more automated and effective in the context of big data. However, PV systems are complex [4], generally including nonlinear [5], [6], uncertain [7], time-correlated [8], multimodal [9], multi-period [10], large-scale [11], or intermittent characteristics [12], resulting in the following problems with the collected data: (1) The characteristic dimension of samples under multi-sensor measurement is relatively high and has high relevance (the data relevance in both decision-making and decision-taking has exponentially increased); (2) Different sampling rates or random loss of data lead to missing observations of some sensors at a specific point in time; (3) Unbalanced data types, such as a limited number of faulty samples under extreme conditions; (4) The data distribution is not uniform: the information from different data sources may be inconsistent. In addition, the PV systems are uncertain with the influence of external disturbances, and the measurements collected are not represented with single values. These problems result in uncertain characteristics of the data and diagnosis spaces and severely limit the representation learning ability of DL algorithms. To improve the safe and stable operation of PV systems, the use of a DL framework for FDD still needs improvement. Generally, the main steps of DL-based FDD are the following: data preprocessing, deep network design, and decision-making. Data preprocessing problem includes (a) the small sample problem. In practical applications, due to problems such as difficulty in data collection and the high cost of sample labeling, the size of the training sample is not large or the amount of data is large but the effective information is insufficient resulting in the problem of small samples in the field of learning. Methods such as transfer learning and generative adversarial networks can be used to solve design problems of deep network-based diagnosis approaches in the case of small samples. (b) Big data storage and analysis: The basis of big data analysis is to extract useful values, suggest conclusions and/or support decision making, and focus on solving problems that cannot be handled within a limited time with existing methods. The priority can be established by several factors such as data preprocessing, fast response speeds, effective reduction of data size, data regularization, relative principal component analysis, etc. In addition, the performance of any diagnosis method depends on the quality of the available process data [13]. Practical measurements usually contain high levels of noise/auto-correlation and are infected with errors that mask

the important features in the data and limit the effectiveness of any process monitoring techniques [14].

Regarding the deep network design, although DL tools have greatly promoted developments in the field of FDD, integrating professional knowledge will help the DL model. Thus, the representation learning of discriminative features can help in reducing the structure of the DL model, and the data regularization processing for specific tasks can help to improve the performance of the FDD. The application of reasonable professional knowledge and prior information helps to reduce the complexity of the monitoring model and improve the diagnosis performance. Besides, the performance of the DL-based FDD relies mainly on the historical data. The rapidity and accuracy of the analysis and effective simplification of the online data to achieve incremental learning of complex dynamic system models and parameter adaptation is a challenging and difficult point. In addition, several issues may have an impact on the diagnosis performance of the DL-based approaches. In general, they are built using default parameters and it is yet to investigate the way the parameter variations affect these approaches. Consequently, DL based on the selection of optimal parameters for FDD must enhance the diagnosis performance.

Faults can be divided into two types depending on their evolution: significant faults and minor faults. The design of a multi-level diagnosis framework in the deep network model will help achieving the real-time monitoring of significant faults and effectively improve the diagnosis of random faults. In addition, significant faults affect the system performance differently, and small faults are also very likely to cause considerable damage. Thus, it is important to develop DL-based algorithms that consider the fusion of faults from different manufacturing environments with different characteristics. Moreover, classical DL algorithms are generally utilized to model the dynamic nature of multivariable PV systems in both the offline training and online updating phases using the updated measurements. Instead, using online extensions of DL models for diagnosis in the first place may reduce the training and update time. In addition, complex PV systems often have problems such as uncertainty, multiple fault occurrence, and fault levels changing with time. If only a single FDD technology is used, the accuracy and generalization will be low. Thus, combining multivariate statistical analysis (such as: PCA [15]–[17], kernel PCA [18]–[20]), signal processing (such as: Fourier transform, multiscale representation [13], [21], interval-valued data representation [22], [23]), and other tools with DL models could improve the performance of the FDD and more specifically decision-making accuracy. It could also reduce the impact of noise, outliers [24], and uncertainties and estimate the severity of the fault location.

The DL-based FDD is mainly divided into three kinds of techniques [25]–[28]: (i) Data preprocessing (DP) → Features Extraction and Selection (FES) → Faults classification (FC) based on DL (FC-DL): This type of method utilizes traditional statistical analysis, signal analysis and other
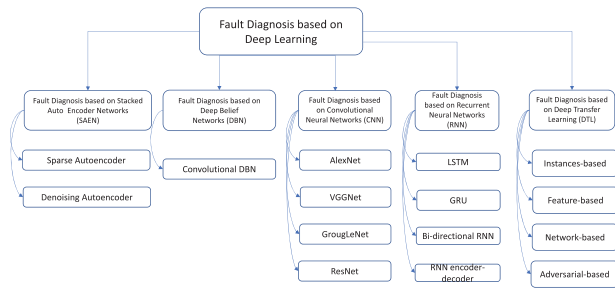
FIGURE 1. Simplified flowchart of fault diagnosis based on deep learning techniques.

methods for data preprocessing and FES, and then applies the DL tool for FC. This allows to reduce the model complexity and improve the diagnosis rate. (ii) DP → FES-DL → FC-DL: In this technique, the DL tool is used for FES, and the FDD is performed based on the extracted high-level features [29], [30], where multi-hidden-layer networks for unsupervised extraction of high-level abstract features are used. It does not require manual intervention or rely on prior knowledge. Combined with multivariate statistical analysis techniques, it is able to provide an efficient diagnosis performance [31]. (iii) DP → (FES and FC)-DL: In this method, the DL tool makes a direct use of the raw-data to perform the FDD. This method belongs to the "End-to End" family, which directly computes the output from the input [32]. The parameters for FES and FC in multi-hidden-layer networks can be optimized collaboratively, and the feature self-learning strategy is adopted to automatically extract the effective features from the large data set to perform the FDD.

This paper mainly discusses the different DL-based FDD techniques for PV systems from the perspective of methodology and five basic architectures: stacked auto encoder network, deep belief network, convolutional neural network, recurrent neural network and deep transfer learning (see Figure 1). The network structure not only determines the effectiveness of feature extraction and selection, but also relates to the complexity of the solution. This paper will explore the research status of these five types of methods respectively, and study their development direction based on the existing problems in the DL-based FDD technology.

The following section presents the most occurring failures in PV systems.

II. COMMON FAILURES IN PV SYSTEMS
In PV systems, the produced PV power depends on various factors such as the nominal characteristics of the components, the power electronics interface, the weather conditions, and failures that may occur in the different stages during the operation (Figure 2).

A. PV MODULE FAILURES
The PV array is the main component of PV systems, where any deficiency associated to the module has a significant
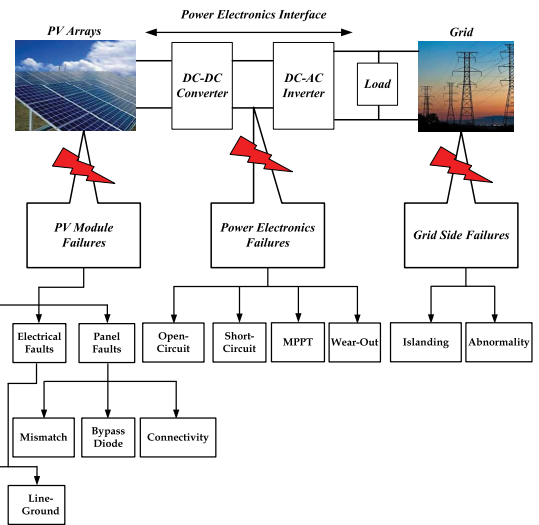

FIGURE 2. Classification of faults in PV systems.

effect on the system performance. In the literature, the following failures have been commonly reported [33]:

1) MISMATCH FAULTS
Grouping non-homogeneous or poorly matched PV cells (non-identical I-V characteristics) causes mismatch faults. This mismatch leads to a reduction of the Maximum Power Point (decreased PV power generation).

2) BYPASS DIODE FAULTS
Usually, a diode connected in anti-parallel to a group of cells (bypass diode) is used in PV modules to prevent cells from shading. Generally, a bypass diode fault is represented by an impedance, short/open circuit, or inverted diode, which causes a mismatch in the I-V characteristics of the cell.

3) CONNECTIVITY FAULTS
Usually, the corrosion/decrease in contact adhesion between two modules lead to a lack of connectivity in PV strings.

4) GROUND FAULTS
A ground fault (GF) in PV modules can be considered as an accidental electrical short-circuit involving ground and one or more current-carrying conductors [40]. The GF may generate DC arcs (and even fire hazards) at the fault point, which raises serious safety concerns.

5) PARTIAL SHADING FAULTS
The operation of PV modules is highly susceptible to partial shading failures, where multiple peaks appear on the P-V characteristics due to the use of a bypass diode.

B. POWER ELECTRONICS INTERFACE FAILURES
Usually, the operation of PV systems is accompanied with failures at different stages. As one of the most important components in PV systems is the power electronics interface,
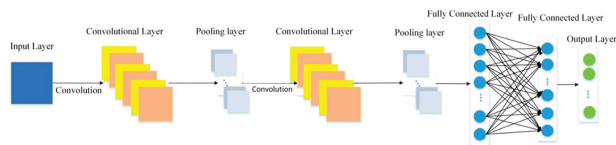
FIGURE 3. Architecture of convolutional neural networks (CNN).

it has been shown that most of the failures are due to the power semiconductor failures.

Many factors may lead to the fatigue of the power electronics components (transistors, diodes). The component fatigue affects mainly the time response and therefore may lead to additional switching losses. Besides, switching of power semiconductors might cause different types of faults.

The three most common power semiconductor failures are the wear-out, open-circuit, and short-circuit faults

C. GRID SIDE FAILURES

Islanding is one of the most important failures to address/detect in grid-connected PV systems. Islanding occurs when a portion of the utility system remains energized while isolated from the grid. This phenomenon can cause safety problems to utility service personnel or related equipment [34].

III. RESEARCH STATUS OF FAULT DIAGNOSIS TECHNOLOGY BASED ON DEEP LEARNING

The DL-based FDD performance is based on the mathematical tool and process models of the plant [35]. The development of deep networks helps to extract high-level and abstract features from the data. When the effective feature representation in the data is relatively extracted, whether it is used for fault classification or regression, better results can be obtained.

A. CONVOLUTIONAL NEURAL NETWORK BASED FAULT DIAGNOSIS

Convolutional Neural Networks (CNN) are built using three types of layers: convolutional layer (CL), pooling layer (PL), and fully connected layer (FCL) [36], [37] (see Figure 3). The CL combines multiple convolution kernels to extract features from the input data or upper layer features, adds matrix element multiplication to the input features in the perceptual field, and adds the deviation [38]. The size of the convolution kernel in the CL controls the extraction of local spatial correlation features in the input information, which can enhance certain features of the original signal while reducing the impact of noise [36]. The PL is responsible for reducing the spatial size of the Convolved feature. It aims to decrease the computational power required to process the data using dimensionality reduction schemes [39]. Furthermore, it is useful for extracting relevant features that are rotationally and positionally invariant, thus maintaining the process of effectively training the model [39]. Adding an FCL is an effective way of learning nonlinear combinations of the high-level features, as represented by the output of the CL. The FCL

is learning a possibly nonlinear function in that space [40]. CNN-based FDD has the following advantages: (i) Industrial system data has multi-source heterogeneity [41]–[44]. The input of CNN can be time series [45]–[47], spectrogram [48], [49], and images [50]–[52], which is suitable for multi-source information processing [41], [53]; (ii) Complex PV systems are often accompanied by random strong magnetic interference, high temperatures. The features extracted by CNN have translation invariance [54], [55], which increases the robustness of the diagnosis algorithm and improves the generalization ability of CNN; (iii) The data that can characterize the faults in PV systems is often submerged in massive real-time data. The generated countermeasure networks can generate samples based on the learning of the probability distribution of real data [56], which is suitable for small sample sizes. The authors in [57]–[60] presented a vector matrix containing statistical characteristics such as the root mean square of the frequency domain signal, the standard deviation, skewness, and kurtosis of the time domain signal as the input to CNN for classification purposes. In addition, in [61], [62], the authors used a Morlet wavelet decomposition tool to obtain the wavelet scale map of the signal, which was used as the input of the CNN for the classification phase. In the developed method, the Rectified Linear Unit (ReLU) was applied as an activation function and used to introduce non-linearity into the network. The work in [63] introduced an adaptive learning rate to construct a hierarchical framework consisting of two CNNs. Therefore, the size of the mode and the adjustment of the adaptive learning rate could promote the algorithm to accelerate the convergence in addition to the prevention of the gradient from disappearing. In addition, given that the traditional linear model cannot capture the complex relationship between sensor data and remaining effective life, the authors in [40], [64] used the time series of multi-channel sensor data. Evidently, one-dimensional (1D) time series can be also used directly as the input of the CNN. For this purpose, the authors in [65] developed a 1D kernel filter to convolve the signal. The proposed method aims to extract high-resolution features for fault detection. The work in [66] used 1D CNN to detect faults by merging feature extraction and post-processing of the raw signal. The authors in [67] presented a comparative study of CNN-based feature learning. The features include raw data, spectrum, and time-frequency data. In [68], a novel full closed-loop -based CNN method for power quality disturbances detection and classification was proposed. The developed approach was able to capture multiscale features and reduce overfitting. To address the problem of small samples, the authors in [69] used prior knowledge to convert normal data into coarse fault data combined with an improved generative adversarial networks (GAN). The proposed technique aims to refine the coarse fault data into data more similar to real faults. The authors in [70] applied a GAN to generate samples with similar distribution to the original signal and utilized a stacked denoising auto-encoder (SDAE) method to pre-train the network, extract fault features, and identify the authenticity of the samples. The developed approach was
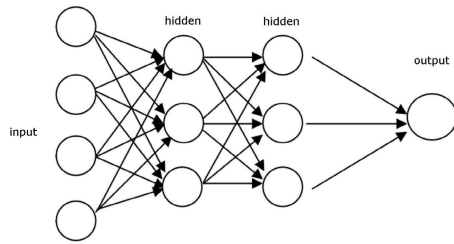
FIGURE 4. Architecture of recurrent neural networks (RNNs).

FIGURE 5. Architecture of stacked auto encoder (SAE) network.

robust to noise and showed a good anti-noise capability in the case of small samples. The work in [71] used a GAN for oversampling the fault operation data to obtain missing fault data. The developed method was proposed to achieve high-precision classification of induction motor faults under different conditions. To repair the "fuzzy" range of the intermediate probability value and enhance the credibility of the reasoning for the fault overlapping area, the authors in [72] applied a GAN to the seismic image in which the feature extraction network was used to extract local and global features from the high-quality image. The reconstruction network then built a high-sensitivity image with a denser sampling rate while retaining the original data and frequency domain information.

B. RECURRENT NEURAL NETWORK BASED FAULT DIAGNOSIS

Recurrent Neural Networks (RNN) are network structures for which the inputs are time-series data and all the nodes are connected in a chain [73]. Unlike multi-layer perceptrons, the RNN have a sense of time and memory of earlier network states allowing them to learn sequences that vary over time [74] (see Figure 4). At present, the most commonly used RNNs are Long Short-Term Memory networks (LSTM) and Gated Recurrent Unit (GRU) networks. By introducing gates, each recurrent unit can adaptively capture the dependence of different time scales to avoid long-term dependence [75]. However, due to gradient exploding and vanishing, there is a length limitation when applying the RNN [76]. Subsequently, RNN variants such as LSTM [76] and GRU [76] neural networks have been developed to deal with long sequence prediction problems. The RNN-based FDD has the following advantages: (1) The inputs of the RNN are time-series data and the depth depends on the length of the input sequence, which is suitable for dynamic PV systems monitoring and prediction; (2) RNN are Turing complete, the chain connection mode is conducive to the extraction and representation of the dynamic nonlinear characteristics of PV systems; (3) The RNN is stable when the length of the learning and testing sequence are different (PV system control is often of variable length and the sampling is irregular). For the time-series signals of PV systems, the authors in [77]–[79] used monotonicity and correlation values to select features as the RNN network inputs, and verified experimentally the
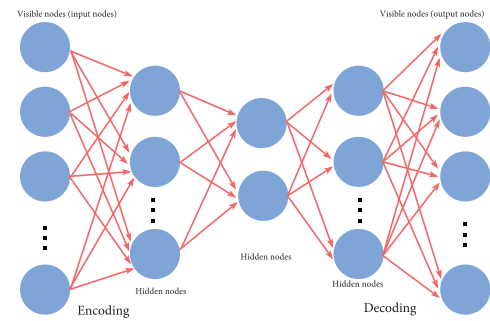
performance of the proposed RNN-based method. The work in [80] proposed an LSTM-based encoder-decoder architecture. The encoder structure converts the input sequence to a fixed-length vector and the decoder structure uses the vector to generate the target sequence and calculate the reconstruction error to use it later for decision making. In the case of multiple faults and large noise, the authors in [81] developed three RNN-based models (vanilla RNN, LSTM, and GRU) for FDD and showed that the LSTM and GRU models outperformed the vanilla RNN. The authors in [82] used GRU in the RNN model as it reduces the parameters by controlling the gate mechanism to alleviate the problem of gradient explosion or disappearance. The authors in [83] proposed sequential FDD based on an LSTM neural network. The developed method can directly classify the raw process data without specific feature extraction and classifier design. It can also adaptively learn the dynamic information in raw data. In [84], the proposed method applies LSTM networks for feature extraction and the selected features are fed into a softmax regression classifier for fault diagnosis. PV system data are characterized by the time relevance in addition to the spatial dependence in the measurement space. For this purpose, the authors in [85] developed a multiscale RNN model for learning both hierarchical and temporal representation. The authors in [86] combined the benefits of both CNN and LSTM to establish a convolutional bidirectional LSTM. In the developed approach, the CNN was used to extract the local features from the original data. Then, a bidirectional LSTM [87] was applied to extract temporal correlation and finally stack a fully connected layer and linear regression layer for predicting the remaining life.

C. STACKED AUTO ENCODER BASED FAULT DIAGNOSIS

Stacked Auto Encoder (SAE) networks are multi-hidden neural networks formed by stacking multiple autoencoder networks [88]. The output of one layer of the auto-encoder network is used as the input of the next layer [89]. Each auto-encoder network consists of two parts: an encoder and a decoder (refer to Figure 5). The encoder converts the network input into the hidden layer representation while the decoder returns the hidden layer representation to the original input [90].

In general, faults tend to occur in high-frequency information corresponding to the higher-order moments of random processes. From the perspective of Taylor expansion, the value of a function in the neighborhood of a point can be represented by infinite series composed of the value of the function at that point and the derivative values of each order. Although the coefficient of the higher-order term is small, it is different from the background details. In these cases, the features are difficult to characterize using traditional methods. The characterization of this information directly affects the performance of the FDD algorithms, especially for small faults that are difficult to detect. As a multi-level network structure model, an SAE network computes higher-order feature representations through multiple nonlinear mappings and expresses more effectively a larger set of functions than shallow networks [91]. FDD based on SAE networks has the following benefits: (i) Most of the data collected from PV systems are 1D signals, and the SAE network structure is simple and suitable for this kind of signal; (ii) In PV applications, data often has unlabeled characteristics and SAE networks are self-learning mechanisms suitable for unsupervised training; (iii) As PV system data contains complex information, the layer-by-layer training method of SAE networks helps to extract the high-order nonlinear features from the data samples and prevents dispersion of the deep network.

Traditional FDD methods are mainly based on the time-frequency analysis of the collected signals [92]. The work in [93] summarizes the traditional feature extraction methods based on frequency domain features. In the auto-encoder network, the depth features obtained using the frequency-domain features, such as low-level inputs, are more suitable for diagnosis systems using Support Vector Machines (SVM) as classifiers. The authors in [94] showed that the features extracted by the stacked denoising auto-encoder network are robust. The performance of the developed technique was assessed by evaluating the impact of the size of the input, the depth of the structure, and the constraint parameters such as sparsity and denoising. Considering that the frequency spectrum reflects the frequency distribution of the data, the authors in [95] introduced the time series frequency spectrum as the input to the SAE network. In [96], a novel SAE-based multiple FDD approach was developed. The proposed technique uses the signal analysis to construct hybrid features and obtain more distinguishing information to overcome the non-stationarity caused by multiple cracks. Finally, the final features are fed as inputs into the SAE network for multiple fault classification. Considering that fault information is mainly reflected in high-frequency mode, the parameters extracted from the previous four modes are used as the network input, which effectively improves the diagnosis performance while simplifying the computation [97]. Given the inadequacy of traditional auto-encoder networks for processing the original input signal, local features, and shifting features, the authors in [98] proposed a local connection network based on a regular sparse auto-encoder. The paper [99] proposed a weight regularization technique to learn weight-invariant facial representations using sparse-stacked denoising autoencoders and deep Boltzmann machines. Due to the limited training data and to prevent data from overfitting, the authors in [100] introduced the ''discard'' technique in the hidden layer of the auto-encoder network. The authors in [101] developed a classification tool using an auto-encoder network. The developed approach applies a prior distribution to the latent space and then uses the mutual information between the sample and predicted distributions for unsupervised clustering. The proposed solution is characterized by its robustness and cross-cutting of the extracted features in a noisy environment. For cross-machine FDD, the work in [102] proposed an approach that holds the potential to largely reduce the expensive labor in data collection for model establishment. Given the process faults under PV multi-modal operation, the authors in [103], [104] introduced the Maximum Square Difference (MSD) to estimate the non-parametric distance between two distributions and proposed a migration learning FDD framework based on sparse auto-encoder networks. The approach developed provided good results, more specifically when the distribution of the testing data was different from that of the training data. The authors in [105] proposed using a time series as the network input to improve the application of time-related information in dynamic PV systems. The work in [106] used a sparse SAE network to limit hidden layer information redundancy, which significantly improves the detection performance of minor faults. The authors in [106] conducted a statistical analysis on the hidden layer features extracted from the SAE network and achieved multi-level FDD based on high-order correlation. The direct use of normal data effectively avoids the imbalance between data categories. To address this issue, the authors in [107] considered the importance of online data diagnosis in the dynamic process and proposed a threshold based on an SAE network. Threshold-adaptive process monitoring technology performs well in diagnosis and reduces the cost and complexity of process modeling.

D. DEEP BELIEF NETWORK BASED FAULT DIAGNOSIS

A Restricted Boltzmann Machine (RBM) is a two-layer neural network composed of a visible layer and a hidden layer. It describes the high-order interaction between variables based on an energy function. The term 'Restricted' means that each edge in the bipartite graph must be connected to one visible unit and one hidden unit [108]. The RBM assumes that when the input data is given, the activation conditions of each hidden unit are independent. Conversely, when the hidden unit state is given, the activation conditions of the visible layer units are independent [109]. The RBM can be a sub-block of a Deep Belief Network (DBN) and a Deep Boltzmann Machine (DBM) (refer to Figure 6). DBN is a multi-hidden-layer probability generation model composed of multiple RBMs and an output layer (usually a classification layer). The joint distribution between observation data and labels is established through layer-by-layer training [110].
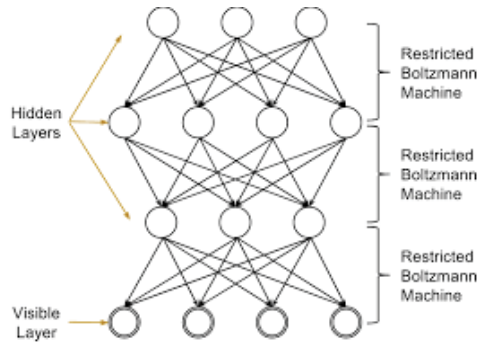
FIGURE 6. Architecture of deep belief network (DBN).

In contrary to the directed/unidirectional connection of the hidden layer in DBN, the DBM is a Boltzmann network with multiple hidden layers. The hidden layer transmits the information and conducts feedback adjustment from top to bottom [110]. Roux and Bengio proved theoretically that as long as the number of hidden units is large enough, the RBM can fit any discrete distribution [109]. FDD based on DBN/DBM has the following advantages: (i) The sample distribution does not necessarily obey restrictive assumptions. RBM uses generative learning to predict the probability distribution of input samples without restrictive assumptions (the PV system has random uncertainty). (ii) RBM expresses data as a probability model through unsupervised learning, which is suitable for sample generative expansion in the case of small samples in PV systems; (iii) The DBN creates activation value sets through feature grouping sequences, which is suitable for simulating and controlling multivariable nonlinear systems (as the PV system control is mostly an unstructured system). The authors in [111], [112] proposed a DBM-based approach considering different characteristics of multi-modal features. The developed approach aims to use the representation learning on time, frequency, and time-frequency domain features, and perform fusion diagnosis in the decision-making phase. In [113], the authors developed an improved convolutional DBN for FDD. First, they used an auto-encoder to compress data and reduce the dimension. Then, the deep model was built with Gaussian visible units to learn the representative features. The presented results showed that the developed strategy provided better accuracy compared to classical DL models. To improve the modeling, [114] added hidden units to the activation function of the sparse SAE network to build a DBN model and to extract the most representative features of the data. In addition, in [115], the authors presented a dual-tree complex wavelet packet method to design the original feature set and constructed an adaptive DBN to improve the network convergence speed and enhance the accuracy of the diagnosis. Paper [116] proposed an improved RBM with a new regularization term to automatically generate features that are suitable for predicting remaining useful life. In [117], the authors proposed a Teager-Kaiser energy operator to estimate the envelope of the instantaneous signal and extract the statistical feature of the data, and then propose

Gaussian-Bernoulli RBM (GRBM) to construct a DBN for real-valued classification [118].

The work in [119] proposed an improved FDD method based on DBN. The so-called DBN developed with a global back-reconstruction (GBR) approach was applied for early crack diagnosis of turbine blades using three-dimensional blade-tip clearance. The authors in [120] proposed a real-time online FDD method that can improve the accuracy of detection, classification, and prediction, while being effective for incipient faults that cannot be detected using statistical tools. A stacked sparse auto-encoder was applied to learn the deep models of fault data and minimize the loss of information. Conversely, [121] merged the benefits of fuzzy Petri nets (FPN) and DBNs to present an adaptive arc generation scheme that presents the label-weight based on confidence-weight to mark the occurrence of a fault. In [122], the authors developed an effective DBN-based approach for detection and diagnosis. An effective DBN model was implemented with an effective distribution of features at each layer of the network to improve the accuracy of the diagnosis at each instant. The authors in [123] proposed an enhanced DBN-based FDD approach that combines the information from multiple sources and enhances the robustness of fault diagnosis.

E. DEEP TRANSFER LEARNING BASED FAULT DIAGNOSIS

The performance of DL-based FDD is closely related to the amount of collected data. To achieve high performance, it is required to generate a large number of samples from the same domain to train the models [124]–[127]. Thereby, when using a large amount of data, DL-based FDD models with complex structures outperform other diagnosis models. Conversely, when a small number of training datasets is generated, the accuracy and reliability of these approaches inevitably decrease [128]. In addition, deep models with a large number of hidden layers can affect the performance of the FDD [129]. Moreover, the training and testing datasets, applied for deep models, have the same feature space and the same distribution [105]. Most statistical models must be rebuilt from scratch using the newly generated training data as the distribution changes. In PV systems, the cost required to collect data again and rebuild the models is very high [130]. Deep transfer learning (DTL) is a promising technique to address these problems [131]. The DTL is a new DL technique that applies existing knowledge to tackle problems in different but related fields, which eases the requirements for data features [132]. The DTL tools can reduce the training time and enhance the classification accuracy by using data in different operating conditions where there is only a small amount of target data [133]–[135]. Recently, it has been widely used for FDD as it can provide accurate results in complex situations where the transfer strategy can help to design a universal diagnostic model [132]. Its main goal is to apply knowledge and skills learned from a data-rich source domain followed by the application to a related target domain with only a small amount of data [133], [136], [137]. The difference between DL and
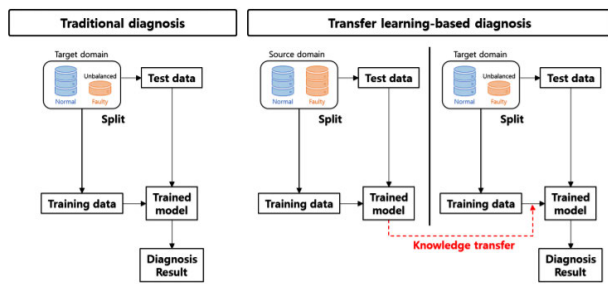
FIGURE 7. Flowchart of deep learning and transfer learning based fault diagnosis.

DTL-based FDD is shown in Figure 7. DL-based FDD aims to split the normal and faulty data into training and testing datasets. Training datasets are applied to train the model for FDD purposes, and then testing datasets are used to measure the performance of the model. The faulty dataset is smaller than that of the normal dataset, which can lead to poor classification performance [128]. In DTL-based FDD, there are two groups of data from different domains: the source and target domains. The source domain is applied to extract knowledge and the target domain uses the extracted knowledge for FDD purposes. In this case, the faulty data in the source field is relatively larger than the one of the target field, which will be exploited for FDD in the target field. The DTL can extract relevant features and perform knowledge transfer between the target and source domains. According to the knowledge transfer, transfer learning techniques can be classified into four categories [133], [138]–[141]: Instances-based DTL [142]–[144], Feature-based DTL [46], [145], [146], Network-based DTL [147]–[149] and Adversarial-based DTL [150]–[153]. Instance-based DTL consists in re-weighting samples from the source field for target field tasks. The Feature-based DTL aims to find the common feature space between the target and source fields. Network-based DTL is based on the assumption that some model parameters can be shared by the source and target fields. Adversarial-based DTL consists in determining the relationship between the samples in the target domain and source domain. Adversarial-based DTL has been proven to provide good results in finding a common latent space between the target and source domains. Thus, it has attracted more attention in the field of transfer learning. DTL has attracted more and more attention in the recent years [154]–[156] and approaches have been applied to several applications such as image recognition [157], text recognition [158], and software defect-recognition [159], as well as FDD. For instance, the authors in [160] developed an adaptive FDD method under different operating conditions to improve the classification accuracy. In [161], the authors developed an automatic DTL based FDD technique for PV systems. The experiments confirmed the very good diagnostic accuracy of the proposed diagnostic method under different simulation conditions. A transfer component analysis approach in [162] was also proposed

for FDD. Despite DTL having been successfully applied for FDD, it still suffers from some limitations. For example, the data used for DTL are all from the same source domain, the knowledge is transferred from just one operating condition to another. However, when multiple related source domains are available, it is difficult to effectively explore general knowledge from those fields and use the information learned in a new related field. Moreover, these methods are not verified with practical data, and the generalization abilities of these methods have not been confirmed. Some previous studies extracted features before applying DTL, which suggests greater requirements for scholars' engineering experience and professional knowledge.

As presented in Table 1, different enhancements were developed in recent years. The table shows improved DL-based FDD techniques and discusses their performance. The proposed techniques focus on the enhancements of basic performance, including reducing the computational complexity, strengthening the robustness of uncertainty and increasing precision. However, in cases of multivariate, uncertain, large and noisy data, it is important to more enhance effectiveness, such as the diagnosis capability, the intelligence level of fault diagnosis, the speed and cost of large data processing, and the integration of statistical and multivariate fault features. Therefore, fault detection and diagnosis more adaptive to the characteristics of data has become a very important topic of research.

IV. EXISTING PROBLEMS AND FUTURE RESEARCH DIRECTIONS

After reviewing the recent literature related to the field of fault detection and diagnosis in PV systems, the following procedure will be adopted.

1) **Enhance Data Preprocessing in Deep Learning**

In PV systems, the data distribution across the different categories is extremely imbalanced. To address the problem of FDD in the case of imbalanced samples, different solutions can be developed, including focal loss function, under-sampling and over-sampling methods at data level, cost-sensitive learning, imbalanced learning, and other models for preprocessing at model level, and then applied as inputs into existing DL-based approaches. Moreover, the larger the size of the training data set is, the lower the effectiveness of the diagnosis is in terms of computation time. This issue limits the implementation of DL methods in practical applications with massive data. To overcome this limitation, improved techniques based on data size reduction frameworks (such as Kmeans metric, Hierarchical K-Means Clustering, Euclidean distance) will be proposed to select the more effective features that can be used as inputs for the DL models for faults classification. In addition, the performance of any diagnosis method depends on the quality of the available process data. The PV measurements usually contain high levels of noise and autocorrelation and are infected

TABLE 1. Summary of DL-based methods for fault diagnosis of PV systems.

Deep Learning Technology	General Features	FDD Technique	Technique Description	Advantages	Reference
Convolutional Neural Network	• Suitable for multisource information processing • Robust diagnosis algorithm and improved generalization ability • Suitable for small sample sizes	Vector matrix (input) + CNN (classifier)	• The input vector (feature representations), formed by the RMS values, standard deviation, skewness, kurtosis, …, is fed to the CNN for classification purposes.	• Good FDD performance in terms of root-mean-square error (RMSE).	[56]
		Morlet wavelet decomposition (input) + CNN (classification)	• The Rectified Linear Unit (ReLU) was applied as an activation function and used to introduce non-linearity into the network	• Capture discriminative information from diverse sets of wavelet coefficients for fault diagnosis. • Effective classification performances evaluated in terms of Sensitivity, Specificity, False alarm rate, Precision and F score.	[61]
		Adaptive learning rate	• Construct a hierarchical framework consisting of two CNNs	• Accelerate the convergence of the algorithm • Prevention of the gradient from disappearing. • Performance evaluated in terms of both fault-pattern recognition and fault-size evaluation.	[63]
		Time series of multi-channel sensor data		• Capture the complex relationship between sensor data and remaining effective life. • FDD performances evaluated in terms of Precision, Recall, F score Accuracy.	[64]
		1D kernel filter	• Convolve the signal	• Extract high-resolution features for FDD. • Performances assessed in terms of average F score.	[65]
		1D CNN	• Merge feature extraction and post-processing of the raw signal	• Avoid any form of conversion • Damage performances evaluated in terms of "Probability of Damage" (PoD).	[66]
		Prior knowledge combined with an improved generative adversarial networks (GAN)	• Use the prior knowledge to convert normal data into coarse fault data	• Address the problem of small samples • Refine the coarse fault data into data more similar to real faults • The average accuracy is used to evaluated the performance.	[67]
		Full closed-loop based CNN		• Capture multiscale features and reduce overfitting. • Effectiveness evaluated in terms of confusion matrix i.e of Sensitivity, Specificity, Precision and Accuracy.	[68]

TABLE 1. *(Continued.)* Summary of DL-based methods for fault diagnosis of PV systems.

		GAN + Stacked Denoising Auto-Encoder (SDAE)	• Apply the GAN to generate samples with similar distribution to the original signal • Apply the SDAE to pre-train the network, extract fault features, and identify the authenticity of the samples	• Robust to noise • Good anti-noise capability in the case of small samples. • Performances evaluated in terms of Loss, Accuracy and execution time.	[69]
Recurrent Neural Network	• The inputs of the RNN are time-series data and the depth depends on the length of the input sequence • RNNs have a sense of time and memory of earlier network states allowing them to learn sequences that vary over time • The RNN is stable when the length of the learning and testing sequence are different	Long Short-Term Memory networks (LSTM)	Gated Recurrent Unit (GRU) networks	• By introducing gates, each recurrent unit can adaptively capture the dependence of different time scales to avoid long-term dependence	• Deal with long sequence prediction problems. • It reduces the parameters by controlling the gate mechanism to alleviate the problem of gradient explosion or disappearance • Avoid long-term dependence • Performances evaluated in terms of prediction accuracy and generalization ability.	[75]
			LSTM-based encoder-decoder architecture	• The encoder structure converts the input sequence to a fixed length vector and the decoder structure uses the vector to generate the target sequence and calculate the reconstruction error to use it later for decision making	• Deal with long sequence prediction problems. • Performances evaluated in terms of Sensitivity, Specificity, Precision, Accuracy and F score.	[80]
			Sequential FDD based on an LSTM neural network		• Deal with long sequence prediction problems. • Can directly classify the raw process data without specific feature extraction and classifier design. • It can also adaptively learn the dynamic information in raw data • Effectiveness evaluated in terms of confusion matrix i.e of Sensitivity, Specificity, Precision and Accuracy. ROC plots are also used for assessment.	[83]

TABLE 1. *(Continued.)* Summary of DL-based methods for fault diagnosis of PV systems.

Stacked Auto Encoder (SAE)		Convolutional bidirectional LSTM	• The CNN is used to extract the local features from the original data and a bidirectional LSTM is applied to extract temporal correlation and finally stack a fully connected layer and linear regression layer for predicting the remaining life	• Deal with long sequence prediction problems. • Combine the benefits of both CNN and LSTM • Comparison among different CPU and GPU platforms. • Performance evaluation in terms of diagnosis accuracy i.e confusion matrix.	[86]
		Vanilla RNN	• Single layer network (with feedback)	• Deal with long sequence prediction problems. • Comparison evaluated in terms of mean absolute error (MAE) and mean squared error (MSE).	[81]
		Multiscale RNN model		• Learning both hierarchical and temporal representation. • Evaluation in terms of log-likelihood averaged over the test sequences	[85]
	• Multi-hidden neural networks formed by stacking multiple autoencoder networks • The output of one layer is used as the input of the next layer • Each auto-encoder network consists of an encoder and a decoder.	Stacked denoising auto-encoder network	• An extension of the basic autoencoder (stochastic version). • Attempt to address identity-function risk by randomly corrupting input (introducing noise) that the autoencoder must then reconstruct (denoise).	• Robust extracted features • The performance is assessed by evaluating the impact of the size of the input, the depth of the structure, and the constraint parameters such as sparsity and denoising.	[94]
		Local connection network	• The network is constructed based on a regular sparse auto-encoder	• Adequacy for processing the original input signal, local features, and shifting features • Diagnosis performances evaluated in terms of accuracy.	[98]
		Sparse stacked denoising autoencoders + deep Boltzmann machines	• A weight regularization using sparse stacked denoising autoencoders and deep Boltzmann machines.	• Learn weight-invariant representations. • Performances evaluated in terms of identification accuracies	[99]
		Maximum Square Difference (MSD) + migration learning FDD framework	• The MSD is proposed to estimate the non-parametric distance between two distributions • The migration learning FDD framework is based on sparse autoencoder networks	• Good results when the distribution of the testing data is different from that of the training data. • Improved classification accuracy i.e confusion marix.	[103]-[104]
		Sparse SAE network	• The proposal conducted a statistical analysis on the hidden layer features extracted from the SAE network and achieved multi-level FDD based on high-order correlation.	• Limit hidden layer information redundancy, which significantly improves the detection performance of minor faults • Testing accuracy of the SAE with different sparsity parameter.	[106]

TABLE 1. *(Continued.)* Summary of DL-based methods for fault diagnosis of PV systems.

				• sparsity parameter. • Visualization of features mined by the proposed method	
Deep Belief Network (DBN)	• The DBN is a multi-hidden layer probability generation model composed of multiple Restricted Boltzmann Machines (RBMs) and an output layer (usually a classification layer). • The RBM can be a sub-block of a DBN and a Deep Boltzmann Machine (DBM). • The joint distribution between observation data and labels is established through layer-by-layer training	DBM-based approach	• Use the representation learning on time, frequency, and time-frequency domain features, and perform fusion diagnosis in the decision-making phase	• Consider different characteristics of multi-modal features • Performances evaluated in terms of classification accuracies	[111]-[112]
		Improved convolutional DBN	• Use an auto-encoder to compress data and reduce the dimension • The deep model is built with Gaussian visible units to learn the representative features	• Better accuracy compared to classical DL models • Diagnosis effectiveness evaluated in terms of average testing accuracy, standard deviation, feature visualization and reconstruction error curves.	[113]
		Enhanced RBM with a new regularization term	• The new regularization term is added to automatically generate features that are suitable for prediction	• Maximize the trendability of the output features, which potentially better represents the degradation pattern of the system. • Performances assessed in terms of prediction accuracy and feature visualization.	[116]
		Teager-Kaiser energy operator + Gaussian-Bernoulli RBM (GRBM)	• The Teager–Kaiser energy operation is proposed to estimate the amplitude envelopes • RBM with Bernoulli hidden units and Gaussian visible units is considered to deal with real-valued signals.	• Reveals the fault patterns contained in the signal (involving transient impacts and noise) • Accurately diagnosing the faults • High reliability • Diagnosis performances evaluated in terms of classification accuracy, mean, standard deviation, and feature visualization.	[117]
		DBN + Global Back-Reconstruction (GBR) mechanism	• The GBR mechanism is realized between the input layer and hidden layers of the DBN by "shortcut connection", and the layer to learn more discriminant features can be determined automatically without prior knowledge	• Optimize the feature learning ability • High diagnostic accuracy • Accuracy and F-score are used to evaluate the experimental results.	[119]
		Enhanced DBN	• Combines the information from multiple sources and enhances the robustness of fault diagnosis	• Fault diagnosis rate (FDR), false positive rate (FPR) and accurate classification rate (ACR) are used for performance evaluation.	[123]

TABLE 1. *(Continued.)* Summary of DL-based methods for fault diagnosis of PV systems.

	• Using data from different domains: the source and target domains • The source domain is applied to extract knowledge and the	Instance-based DTL	• Applies existing knowledge to tackle problems in different but related fields, which eases the requirements for data features.	• Transfer diagnosis cases of fault severities and compound faults are used for verification.	[142]
Deep Transfer Learning (DTL)	target domain uses the extracted knowledge for FDD purposes • The faulty data in the source field is relatively larger than the one of the target field • Can extract relevant features and perform knowledge transfer between the target and source domains	Feature-based DTL	• Re-weighting samples from the source field for target field tasks	• Boosting the quality of deep learning models • Good diagnosis accuracies in terms of confusion matrix, validation loss and feature visualization.	[144]
			• Learns features from source data and adjusts the parameters of neural networks accordingly. Some parameters are transferred from source task to target task. Small amount of target data is considered.	• Performances evaluated in terms of classification accuracy and training time.	[145]
			• In feature extraction, sparse filtering with HKL divergence is proposed to learn sharing and discriminative features of the source and target domains. In feature classification, HKL divergence is introduced into softmax regression to link the domain adaptation with health conditions.	• Good diagnosis accuracies in terms of average accuracy and confusion matrix.	[146]
		Network-based DTL	• Combines convolutional neural network and multiple domain adaptation techniques. • Uses computer simulations to overcome data scarcity in process fault diagnosis. • Uses domain adaptation to reduce the mismatch between the simulation and physical domains.	• Diagnosis performances evaluated in terms of diagnosis accuracy and associated standard deviation	[147]-[149]
		Adversarial-based DTL	• It includes three training phases: regular training, maximum discrepancy training and minimum discrepancy training. These three phases are adversarial to each other to adjust the model to be more adaptive.	• Provides good results in finding a common latent space between the target and source domains. • Diagnosis effectiveness assessed in terms of prediction accuracy.	[150]-[153]

with errors that mask the important features in the data and limit the effectiveness of any process monitoring techniques. Therefore, multiscale data representation is a forceful data analysis tool that decomposes the original process samples into multiscale components to provide an effective separation of the deterministic and stochastic features from the data. The data is decomposed at multiple scales using low-pass and high-pass filters and the noise is separated from the important characteristics. Thus, the implementation of diagnosis frameworks that unify DL methods and multiscale representation schemes may improve the performance of classical FDD using DL approaches. Therefore, multiscale DL methods combining the advantages of the former with those of multiscale representation should improve the diagnosis results. The developed techniques will provide the grid operators and power system designers with significant information to design an optimal solar PV plant, as well as to manage power supply and demand. This task aims to develop various data preprocessing methods in terms of characteristics and performance. The developed techniques will be used due to their ability to solve non-linear, dynamic, and multivariate data structures of PV systems.

- **Enhance Data Preprocessing using Multiscale Representation**
  Data preprocessing approaches are completely based on the process data. Therefore, the quality of the data plays an important role in the accuracy of the derived features. It is known that the measured process data are contaminated with noise that degrades their usefulness in the diagnosis. Therefore, the measurement noise needs to be filtered to enhance the quality of the extracted features. Multiscale data representation is a powerful data analysis tool that has been effectively used to enhance the quality of various process data preprocessing methods [14], [163], [164]. In addition, multiscale filtering can also be used to further enhance the data preprocessing accuracy by developing multiscale data preprocessing techniques. The developed techniques will be utilized to enhance the monitoring and diagnosis taking into account the measurement noise and the dynamics of the PV systems. Multiscale filtering is an effective data filtering method. However, pre-filtering the data before constructing the models may not provide the desired advantages of multiscale filtering. This is due to the fact that the data pre-filtering may eliminate some features in the data that are important to the model [165]. Therefore, data preprocessing and multiscale filtering need to be integrated to achieve the desired model accuracy. One way to do that is to filter the data using multiscale filtering at different decomposition depths,

construct data preprocessing methods using the filtered data from each decomposition depth, and then select the model that provides the optimum prediction accuracy. The proposed multiscale data preprocessing technique will combine the advantages of both multiscale estimation and data preprocessing.

- **Develop Data Preprocessing Methods for Uncertain PV Systems using Interval-Valued Data Representation and Dimensionality Reduction**
  New interval data preprocessing approaches can be developed to deal with uncertainties in PV systems. In fact, real PV systems are often affected by different types of uncertainties, mainly due to the measurement errors and noise, as well as current and voltage variability. The uncertainty in the model may be addressed by considering the interval-valued data. The developed techniques will enhance the above-proposed data preprocessing approaches by taking into account the irradiance, current, voltage, and temperature uncertainties.

Neither measurements nor estimations are 100 % accurate, so in reality the actual value $x_j^*(k)$ of a variable can deviate from the measured one $x_j^c(k)$. The measurement errors are defined as $\delta x_j(k) = x_j^c(k) - x_j^*(k)$. Usually, the sensors manufacturers provide an upper bound $\delta x_j(k)$ on the measurement error. Hence, once a measurement $x_j^c(k)$ is available, one should know that the actual (unknown) value $x_j^*(k)$ of the measured variable belongs to the interval $x_j^*(k) = [x_j^-(k) \ x_j^+(k)]$, where $x_j^-(k) = x_j^c(k) - \delta x_j(k)$ and $x_j^+(k) = x_j^c(k) + \delta x_j(k)$.

An interval valued data $[x(k)]$ refers to a set of numbers enclosed in an interval on the real line, usually expressed as $[x] = [x^-(k) \ x^+(k)]$, where $x^-(k), x^+(k) \in \mathcal{R}$ and $x^-(k) \le x^+(k)$.

We will start by describing the properties of the interval-valued variables [166]. An interval-valued variable $[X_j] \subset \mathcal{R}$ is represented by a series of sets of values delimited by ordered couples of bounds referred as minimum and maximum: $[X_j] = \{[x_j(1)], [x_j(2)], \ldots, [x_j(n)]\}$, where $[x_j(k)] \equiv [x^-_j(k), x^+_j(k)] \ \forall \ k \in \{1, \ldots, n\}$ and $x_j^-(k) \le x_j^+(k)$. The generic interval $[x_j(k)]$ can be also expressed by a couple $\{x_j^c(k), x_j^r(k)\}$ and that this is a bi-univocal relationship, where:

$$x_j^c(k) = \frac{1}{2}(x_j^+(k) + x_j^-(k)), \qquad (1)$$

and

$$x_j^r(k) = \frac{1}{2}(x_j^+(k) - x_j^-(k)). \qquad (2)$$

The interval-valued data matrix $[X]$ is an $n \times m$ data matrix, given by,

$$[X] = \begin{pmatrix} [x_1^-(1), x_1^+(1)] & \cdot & \cdot & [x_m^-(1), x_m^+(1)] \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ [x_1^-(n), x_1^+(n)] & \cdot & \cdot & [x_m^-(n), x_m^+(n)] \end{pmatrix}, \tag{3}$$

where $x_j^-(k) \leq x_j^+(k)$ for all $k = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$.

Two Euclidean Distance (ED)-based interval-valued data can be developed to remove the irrelevant and redundant samples during the data preprocessing task.

The use of interval-valued data is motivated by the need of size reduction of massive datasets in some applications. An interval-valued variable $[x_{j,k}]$, can be determined using a lower and upper bound [167], such as $[x_j(k)] = [\underline{x}_{j,k}, \overline{x}_{j,k}]$, where $k \in \{1, \ldots, N\}$, $\underline{x}_{j,k} \leq \overline{x}_{j,k}$, and N is the number of samples.

Given an $N \times m$ classical training data matrix X, where m is the number of variables and N is the number of samples, the interval data matrix $[X]$ can be constructed as per:

$$[X] = \begin{pmatrix} [\underline{x}_{1,1}, \overline{x}_{1,1}] & \cdot & \cdot & [\underline{x}_{1,m}, \overline{x}_{1,m}] \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ [\underline{x}_{N1}, \overline{x}_{N,1}] & \cdot & \cdot & [\underline{x}_{N,m}, \overline{x}_{N,m}] \end{pmatrix} \tag{4}$$

where, the lower X^L and upper X^U bound matrices are respectively defined by:

$$X^L = \begin{pmatrix} \underline{x}_{1,1} & \cdot & \cdot & \underline{x}_{1,m} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \underline{x}_{N,1} & \cdot & \cdot & \underline{x}_{N,m} \end{pmatrix} \tag{5}$$

$$X^U = \begin{pmatrix} \overline{x}_{1,1} & \cdot & \cdot & \overline{x}_{1,m} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \overline{x}_{N,1} & \cdot & \cdot & \overline{x}_{N,m} \end{pmatrix} \tag{6}$$

The interval-valued variable $[x_{j,k}]$ can be also expressed by a couple $\{x_{j,k}^c, x_{j,k}^r\}$.

The center $x_{j,k}^c$ of the interval is given by:

$$x_{j,k}^c = \frac{1}{2}(\overline{x}_{j,k} + \underline{x}_{j,k}) \tag{7}$$

and the range $x_j^r(k)$ of the interval is defined by:

$$x_j^r(k) = \frac{1}{2}(\overline{x}_{j,k} - \underline{x}_{j,k}) \tag{8}$$

In this case, the center and range matrices are respectively defined by:

$$X^c = \frac{1}{2} \begin{pmatrix} \underline{x}_{11} + \overline{x}_{1,1} & \cdot & \cdot & \underline{x}_{1m} + \overline{x}_{1,m} \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ \underline{x}_{N,1} + \overline{x}_{N,1} & \cdot & \cdot & \underline{x}_{N,m} + \overline{x}_{N,m} \end{pmatrix} \tag{9}$$

$$X^r = \frac{1}{2} \begin{pmatrix} \underline{x}_{11} - \overline{x}_{1,1} & \cdot & \cdot & \underline{x}_{1,m} - \overline{x}_{1,m} \\ & \cdot & & \cdot \\ & \cdot & & \cdot \\ \underline{x}_{N,1} - \overline{x}_{N,1} & \cdot & \cdot & \underline{x}_{N,m} - \overline{x}_{N,m} \end{pmatrix} \tag{10}$$

By the concatenation of the center and range matrices, the new data matrix $\mathbf{X}_{CR}$ can expressed by:

$$\mathbf{X}_{cr} = \begin{bmatrix} \mathbf{X}^c & \mathbf{X}^r \end{bmatrix} \tag{11}$$

2) **Enhance Deep Network Design in Deep Learning**

The representation learning of features can help in reducing the structure of the DL model, and the data regularization processing for specific tasks can improve the performance of FDD. The application of reasonable professional knowledge and prior information helps to reduce the complexity of the process monitoring model and improve the diagnosis performance. Moreover, the performance of DL-based FDD mainly relies on the historical data. Although historical data contains the operating mechanism of complex systems, PV systems are dynamic production processes and the latest changes in the current operating state also include the cumulative relevance of the production process. The rapid/accurate analysis and effective simplification of the online data to achieve the incremental learning of complex dynamic system models and parameter adaptation is in fact a challenging and difficult task. In addition, several challenges may have an impact on the FDD results using DL-based approaches. In general, the DL approaches based FDD are built using default parameters and it is yet to investigate how the parameter variations affect these approaches. Consequently, the DL based on optimal parameter selection for diagnosing faults can be developed. The parameters to be optimized include the number of hidden layer nodes and the activation function for extracting features and reconstructing inputs. This task reduces the requirements for research experience during parameter tuning and avoids the need for tedious manual tuning. Moreover, the optimized DL model can achieve improved diagnosis performance. The optimization tools including Orca Optimization Algorithm (OOA), Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and Multi-Objective Optimization (MOO) will be employed to optimize the DL parameters.

3) **Enhance Decision Making in Deep Learning**

In PV systems, the significant faults affect considerably the system performance. However, small faults are also very likely to cause a considerable damage. Thus, DL-based algorithms, merging faults with different characteristics, can be developed. Moreover, classical DL algorithms are generally utilized to model the dynamic nature of multivariable PV systems in both the offline training and online updating phases using the newly arrived measurements. Instead, using the online extensions of DL models for diagnosis may reduce the training and update time. In addition,

complex PV systems often have problems such as uncertainty and multiple fault concurrency. If only a single FDD technique is used, the accuracy and generalization will be low. Thus, combining multivariate statistical analyses (such as PCA and kernel PCA), signal processing (such as Fourier transform, multiscale representation, interval-valued data representation), and other tools with DL models could improve the performance of FDD and more specifically the decision-making accuracy. It could also reduce the impact of noise, outliers, and uncertainties, and estimate the severity of the fault location. It has been shown in [168] that the introduction of selected features using PCA in the DL classifiers (i.e. NN and RNN) enhances the classification accuracy compared to conventional raw data-based classifiers. However, in the analysis, we assumed that the features were extracted and selected using a linear PCA, and the PV faults were classified using classical DL classifiers. There are different ways to enhance the performance of the techniques developed in [27]. To solve the main problem of the linear characteristics of the PCA in high-dimensional spaces, the kernel PCA (KPCA) will be applied to extract high-order statistical information in the DL parameter space. Although KPCA can extract nonlinear features in high-dimensional spaces, it increases the space and time complexity compared to PCA. To improve the use of KPCA, reduced extensions will be proposed. The reduced KPCA (RKPCA) uses different dimension reduction metrics (such as Kmeans metric, Hierarchical K-Means Clustering, Euclidean distance) so that only the effective samples are selected and applied to build the KPCA model. To address the problem of uncertainties in the PV systems, interval RKPCA models will be also proposed.

4) **Develop Enhanced Multiple Deep Learners**

The last task in FDD is to build a learner model. For this purpose, an enhanced DL technique can be developed. The developed technique that merges different learners should improve the diagnosis performance. In this task, a novel design optimization technique based on multiple deep learners, which includes the above improvements (Tasks 1-3), can be developed. In this technique, a hybrid method incorporating multiple deep learners that fits high-level information will be deployed. However, the diagnosis accuracy is based on the weighting factors linking the deep learners. Thus, an optimal selection procedure of the weighting factors is required to further improve the FDD performance of the deep learner. The optimization problem is addressed so that the miss-classification and execution time are jointly minimized. Therefore, an enhanced multiple deep learner method will be proposed to obtain better diagnosis ability compared to classical standalone deep learners. The developed technique will contribute to the reduction of the overall diagnosis error

and will have the ability to combine various models. To do that, multivariate and dynamic features will be considered in designing multiple learning models. Classical multiple models ignore the time-dependence of PV measurements. However, the PV system data are frequently time-correlated. Accordingly, the dynamic and multivariate nature of the measurements will be considered when designing the prediction models by using multivariate and dynamic techniques (such as PCA, kernel PCA, and Dynamic kernel PCA).

V. CONCLUSION

Data-based fault detection and diagnosis (FDD) is an effective solution towards high performance and reliability PV systems. The most well-known data-driven methods are deep learning (DL) approaches. Therefore, this paper discussed the DL-based FDD in PV systems. In the present review paper, the DL-based FDD have been classified into five categories: FDD based on convolutional neural network (CNN), FDD based on recurrent neural network (RNN), FDD based on stacked auto encoder network (SAEN), FDD based on deep belief network (DBN) and FDD based on deep transfer learning (DTL), where their main advantages and drawbacks were indicated. Finally, the topic has been studied at three levels including data preprocessing, deep network design, and decision-making module. Furthermore, other FDD solutions have been proposed by considering uncertainties, complexity, multivariate and dynamic natures of industrial systems. The biggest advantage of the DL-based FDD algorithms is their capability to learn high-level features from data in a high-order, non-linear, and adaptive manners. Because of this powerful feature representation learning effectiveness, intelligent FDD techniques become more effective. Although DL-based FDD has greatly promoted the development of the diagnosis research field, it is relatively considered as a new concept and further in-depth investigations are required.

REFERENCES

[1] M. Mansouri, A. Al-Khazraji, M. Hajji, M. F. Harkat, H. Nounou, and M. Nounou, "Wavelet optimized EWMA for fault detection and application to photovoltaic systems," *Sol. Energy*, vol. 167, pp. 125–136, Jun. 2018.

[2] A. Haque, K. V. S. Bharath, M. A. Khan, I. Khan, and Z. A. Jaffery, "Fault diagnosis of photovoltaic modules," *Energy Sci. Eng.*, vol. 7, no. 3, pp. 622–644, Mar. 2019.

[3] V. S. B. Kurukuru, F. Blaabjerg, M. A. Khan, and A. Haque, "A novel fault classification approach for photovoltaic systems," *Energies*, vol. 13, no. 2, p. 308, Jan. 2020.

[4] A. Sayed, M. El-Shimy, M. El-Metwally, and M. Elshahed, "Reliability, availability and maintainability analysis for grid-connected solar photovoltaic systems," *Energies*, vol. 12, no. 7, p. 1213, Mar. 2019.

[5] K. Bataineh and N. Eid, "A hybrid maximum power point tracking method for photovoltaic systems for dynamic weather conditions," *Resources*, vol. 7, no. 4, p. 68, Nov. 2018.

[6] V. S. B. Kurukuru, A. Haque, A. K. Tripathi, and M. A. Khan, "Condition monitoring of IGBT modules using online TSEPs and data-driven approach," *Int. Trans. Electr. Energy Syst.*, vol. 31, no. 8, Aug. 2021, Art. no. e12969.

[7] K. Dhibi, R. Fezai, M. Mansouri, A. Kouadri, M.-F. Harkat, K. Bouzara, H. Nounou, and M. Nounou, "A hybrid approach for process monitoring: Improving data-driven methodologies with dataset size reduction and interval-valued representation," *IEEE Sensors J.*, vol. 20, no. 17, pp. 10228–10239, Sep. 2020.

[8] T. AlSkaif, S. Dev, L. Visser, M. Hossari, and W. van Sark, "A systematic analysis of meteorological variables for PV output power estimation," *Renew. Energy*, vol. 153, pp. 12–22, Jun. 2020.

[9] Y. Zhao, Q. Liu, D. Li, D. Kang, Q. Lv, and L. Shang, "Hierarchical anomaly detection and multimodal classification in large-scale photovoltaic systems," *IEEE Trans. Sustain. Energy*, vol. 10, no. 3, pp. 1351–1361, Jul. 2019.

[10] H. Karunathilake, K. Hewage, T. Prabatha, R. Ruparathna, and R. Sadiq, "Project deployment strategies for community renewable energy: A dynamic multi-period planning approach," *Renew. Energy*, vol. 152, pp. 237–258, Jun. 2020.

[11] X. Li, Q. Yang, Z. Lou, and W. Yan, "Deep learning based module defect analysis for large-scale photovoltaic farms," *IEEE Trans. Energy Convers.*, vol. 34, no. 1, pp. 520–529, Mar. 2018.

[12] J. Li, S. Chen, Y. Wu, Q. Wang, X. Liu, L. Qi, X. Lu, and L. Gao, "How to make better use of intermittent and variable energy? A review of wind and photovoltaic power consumption in China," *Renew. Sustain. Energy Rev.*, vol. 137, Mar. 2021, Art. no. 110626.

[13] G. Jiang, H. He, J. Yan, and P. Xie, "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," *IEEE Trans. Ind. Electron.*, vol. 66, no. 4, pp. 3196–3207, Apr. 2018.

[14] M. Mansouri, M. N. Nounou, and H. N. Nounou, "Multiscale kernel PLS-based exponentially weighted-GLRT and its application to fault detection," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 3, no. 1, pp. 49–58, Nov. 2017.

[15] L. I. Kuncheva and W. J. Faithfull, "PCA feature extraction for change detection in multidimensional unlabeled data," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 1, pp. 69–80, Jan. 2013.

[16] S. Li and J. Wen, "A model-based fault detection and diagnostic methodology based on PCA method and wavelet transform," *Energy Buildings*, vol. 68, pp. 63–71, Jan. 2014.

[17] T. Rato, M. Reis, E. Schmitt, M. Hubert, and B. De Ketelaere, "A systematic comparison of PCA-based statistical process monitoring methods for high-dimensional, time-dependent processes," *AIChE J.*, vol. 62, no. 5, pp. 1478–1493, May 2016.

[18] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Comput.*, vol. 10, no. 5, pp. 1299–1319, Jul. 1998.

[19] J.-M. Lee, C. K. Yoo, S. W. Choi, P. A. Vanrolleghem, and I.-B. Lee, "Nonlinear process monitoring using kernel principal component analysis," *Chem. Eng. Sci.*, vol. 59, no. 1, pp. 223–234, Jan. 2004.

[20] J. H. Cho, J. M. Lee, S. W. Choi, D. Lee, and I. B. Lee, "Fault identification for process monitoring using kernel principal component analysis," *Chem. Eng. Sci.*, vol. 60, no. 1, pp. 279–288, 2005.

[21] Y. Zhang and C. Ma, "Fault diagnosis of nonlinear processes using multiscale KPCA and multiscale KPLS," *Chem. Eng. Sci.*, vol. 66, no. 1, pp. 64–72, 2011.

[22] T. Ait-Izem, M.-F. Harkat, M. Djeghaba, and F. Kratz, "Sensor fault detection based on principal component analysis for interval-valued data," *Qual. Eng.*, vol. 30, no. 4, pp. 1–13, 2017.

[23] F. D. A. de Carvalho, E. D. A. L. Neto, and C. P. Tenorio, "A new method to fit a linear regression model for interval-valued data," in *Proc. Annu. Conf. Artif. Intell.* Berlin, Germany: Springer, 2004, pp. 295–306.

[24] M. Kang, M. R. Islam, J. Kim, J. M. Kim, and M. Pecht, "A hybrid feature selection scheme for reducing diagnostic performance deterioration caused by outliers in data-driven diagnostics," *IEEE Trans. Ind. Electron.*, vol. 60, no. 5, pp. 3299–3310, May 2016.

[25] M. Mansouri, R. Fezai, M. Trabelsi, M. Hajji, M.-F. Harkat, H. Nounou, M. N. Nounou, and K. Bouzrara, "A novel fault diagnosis of uncertain systems based on interval Gaussian process regression: Application to wind energy conversion systems," *IEEE Access*, vol. 8, pp. 219672–219679, 2020.

[26] R. Fezai, K. Dhibi, M. Mansouri, M. Trabelsi, M. Hajji, K. Bouzrara, H. Nounou, and M. Nounou, "Effective random forest-based fault detection and diagnosis for wind energy conversion systems," *IEEE Sensors J.*, vol. 21, no. 5, pp. 6914–6921, Mar. 2021.

[27] S. Gharsellaoui, M. Mansouri, M. Trabelsi, M.-F. Harkat, S. S. Refaat, and H. Messaoud, "Interval-valued features based machine learning technique for fault detection and diagnosis of uncertain HVAC systems," *IEEE Access*, vol. 8, pp. 171892–171902, 2020.

[28] K. Dhibi, R. Fezai, M. Mansouri, M. Trabelsi, A. Kouadri, K. Bouzara, H. Nounou, and M. Nounou, "Reduced kernel random forest technique for fault detection and classification in grid-tied PV systems," *IEEE J. Photovolt.*, vol. 10, no. 6, pp. 1864–1871, Nov. 2020.

[29] R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A comprehensive review of dimensionality reduction techniques for feature selection and feature extraction," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 56–70, May 2020.

[30] K. E. Pilario, M. Shafiee, Y. Cao, L. Lao, and S.-H. Yang, "A review of kernel methods for feature extraction in nonlinear process monitoring," *Processes*, vol. 8, no. 1, p. 24, Dec. 2019.

[31] J. Kim and N. Moon, "A deep bidirectional similarity learning model using dimensional reduction for multivariate time series clustering," *Multimedia Tools Appl.*, pp. 1–13, Jan. 2021, doi: 10.1007/s11042-020-10476-6.

[32] J. Xie, G. Du, C. Shen, N. Chen, L. Chen, and Z. Zhu, "An end-to-end model based on improved adaptive deep belief network and its application to bearing fault diagnosis," *IEEE Access*, vol. 6, pp. 63584–63596, 2018.

[33] R. Fazai, K. Abodayeh, M. Mansouri, M. Trabelsi, H. Nounou, M. Nounou, and G. E. Georghiou, "Machine learning-based statistical testing hypothesis for fault detection in photovoltaic systems," *Sol. Energy*, vol. 190, pp. 405–413, Sep. 2019.

[34] M. Trabelsi and H. Abu-Rub, "A unique active anti-islanding protection for a quasi-Z-source based power conditioning system," in *Proc. IEEE Appl. Power Electron. Conf. Expo. (APEC)*, Mar. 2015, pp. 2237–2243.

[35] M. S. Mirnaghi and F. Haghighat, "Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review," *Energy Buildings*, vol. 229, Dec. 2020, Art. no. 110492.

[36] X. Wang, Y. Zhao, and F. Pourpanah, "Recent advances in deep learning," Tech. Rep., 2020, doi: 10.13140/RG.2.2.24831.10403.

[37] M. Manohar, E. Koley, S. Ghosh, D. K. Mohanta, and R. C. Bansal, "Spatio-temporal information based protection scheme for PV integrated microgrid under solar irradiance intermittency using deep convolutional neural network," *Int. J. Electr. Power Energy Syst.*, vol. 116, Mar. 2020, Art. no. 105576.

[38] F. Aziz, A. U. Haq, S. Ahmad, Y. Mahmoud, M. Jalal, and U. Ali, "A novel convolutional neural network-based approach for fault classification in photovoltaic arrays," *IEEE Access*, vol. 8, pp. 41889–41904, 2020.

[39] 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.

[40] X. Lu, P. Lin, S. Cheng, Y. Lin, Z. Chen, L. Wu, and Q. Zheng, "Fault diagnosis for photovoltaic array based on convolutional neural network and electrical time series graph," *Energy Convers. Manage.*, vol. 196, pp. 950–965, Sep. 2019.

[41] A. I. Khalyasmaa, S. A. Eroshenko, V. A. Tashchilin, H. Ramachandran, T. Piepur Chakravarthi, and D. N. Butusov, "Industry experience of developing day-ahead photovoltaic plant forecasting system based on machine learning," *Remote Sens.*, vol. 12, no. 20, p. 3420, Oct. 2020.

[42] Y. Ma, L. Zhao, Y. Zhang, Q. Zhou, M. Zhang, and S. Yang, "Multi-objective optimal scheduling of power systems based on complementary characteristics of heterogeneous energy sources," in *Proc. IEEE 4th Conf. Energy Internet Energy Syst. Integr. (EI)*, Oct. 2020, pp. 1533–1538.

[43] A. Gaabour, A. Metatla, R. Kelaiaia, F. Bourennani, and A. Kerboua, "Recent bibliography on the optimization of multi-source energy systems," *Arch. Comput. Methods Eng.*, vol. 26, no. 4, pp. 809–830, Sep. 2019.

[44] P. Dagnely, T. Ruette, T. Tourwé, and E. Tsiporkova, "Ontology-driven multilevel sequential pattern mining: Mining for gold in event logs of photovoltaic plants," in *Proc. Int. Conf. Intell. Syst. (IS)*, Sep. 2018, pp. 1–7.

[45] M. Canizo, I. Triguero, A. Conde, and E. Onieva, "Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study," *Neurocomputing*, vol. 363, pp. 246–260, Oct. 2019.

[46] R. Ye and Q. Dai, "Implementing transfer learning across different datasets for time series forecasting," *Pattern Recognit.*, vol. 109, Jan. 2021, Art. no. 107617.

[47] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *J. Syst. Eng. Electron.*, vol. 28, no. 1, pp. 162–169, Feb. 2017.

[48] G. Kovács, L. Tóth, D. Van Compernolle, and S. Ganapathy, "Increasing the robustness of CNN acoustic models using autoregressive moving average spectrogram features and channel dropout," *Pattern Recognit. Lett.*, vol. 100, pp. 44–50, Dec. 2017.

[49] G. C. Jana, R. Sharma, and A. Agrawal, "A 1D-CNN-spectrogram based approach for seizure detection from EEG signal," *Proc. Comput. Sci.*, vol. 167, pp. 403–412, Jan. 2020.

[50] Z. Lin, K. Ji, X. Leng, and G. Kuang, "Squeeze and excitation rank faster R-CNN for ship detection in SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 5, pp. 751–755, May 2018.

[51] Y. Li, H. Su, C. R. Qi, N. Fish, D. Cohen-Or, and L. J. Guibas, "Joint embeddings of shapes and images via CNN image purification," *ACM Trans. Graph.*, vol. 34, no. 6, pp. 1–12, Nov. 2015.

[52] A. R. Beeravolu, S. Azam, M. Jonkman, B. Shanmugam, K. Kannoorpatti, and A. Anwar, "Preprocessing of breast cancer images to create datasets for deep-CNN," *IEEE Access*, vol. 9, pp. 33438–33463, 2021.

[53] L. Shuai, L. Yuanning, Z. Xiaodong, H. Guang, C. Jingwei, Z. Qixian, W. Zukang, L. Xinlong, and W. Chaoqun, "Multi-source feature fusion and entropy feature lightweight neural network for constrained multi-state heterogeneous iris recognition," *IEEE Access*, vol. 8, pp. 53321–53345, 2020.

[54] B. Lin, S. Deng, H. Gao, and J. Yin, "A multi-scale activity transition network for data translation in EEG signals decoding," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, early access, Sep. 15, 2020, doi: 10.1109/TCBB.2020.3024228.

[55] E. Kauderer-Abrams, "Quantifying translation-invariance in convolutional neural networks," 2017, *arXiv:1801.01450*. [Online]. Available: http://arxiv.org/abs/1801.01450

[56] S. M. Razak and B. Jafarpour, "Convolutional neural networks (CNN) for feature-based model calibration under uncertain geologic scenarios," *Comput. Geosci.*, vol. 24, no. 4, pp. 1625–1649, Aug. 2020.

[57] R. Pierdicca, E. S. Malinverni, F. Piccinini, M. Paolanti, A. Felicetti, and P. Zingaretti, "Deep convolutional neural network for automatic detection of damaged photovoltaic cells," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vols. XLII–2, pp. 893–900, May 2018.

[58] R. Al-Mashhadani, G. Alkawsi, Y. Baashar, A. A. Alkahtani, F. H. Nordin, and W. Hashim, "Deep learning methods for solar fault detection and classification: A review," *Inf. Sci. Lett.*, vol. 10, no. 2, p. 13, 2021.

[59] S. N. Venkatesh and V. Sugumaran, "Fault detection in aerial images of photovoltaic modules based on deep learning," in *Proc. IOP Conf., Mater. Sci. Eng.*, 2021, vol. 1012, no. 1, Art. no. 012030.

[60] B. Basnet, H. Chun, and J. Bang, "An intelligent fault detection model for fault detection in photovoltaic systems," *J. Sensors*, vol. 2020, pp. 1–11, Jun. 2020.

[61] M. H. Zhao, M. Kang, B. Tang, and M. Pecht, "Multiple wavelet coefficients fusion in deep residual networks for fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 66, no. 6, pp. 4696–4706, Jun. 2018.

[62] T. Li, Z. Zhao, C. Sun, L. Cheng, X. Chen, R. Yan, and R. X. Gao, "WaveletKernelNet: An interpretable deep neural network for industrial intelligent diagnosis," *IEEE Trans. Syst., Man, Cybern. Syst.*, early access, Jan. 20, 2021, doi: 10.1109/TSMC.2020.3048950.

[63] X. Guo, L. Chen, and C. Shen, "Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis," *Measurement*, vol. 93, pp. 490–502, Nov. 2016.

[64] S. Samara and E. Natsheh, "Intelligent PV panels fault diagnosis method based on NARX network and linguistic fuzzy rule-based systems," *Sustainability*, vol. 12, no. 5, p. 2011, Mar. 2020.

[65] C.-H. Hsieh, Y.-S. Li, B.-J. Hwang, and C.-H. Hsiao, "Detection of atrial fibrillation using 1D convolutional neural network," *Sensors*, vol. 20, no. 7, p. 2136, Apr. 2020.

[66] N. V. Sridharan and V. Sugumaran, "Convolutional neural network based automatic detection of visible faults in a photovoltaic module," *Energy Sour. A, Recovery, Utilization, Environ. Effects*, pp. 1–16, Mar. 2021, doi: 10.1080/15567036-021-190575-3.

[67] T. Han, D. Jiang, Y. Sun, N. Wang, and Y. Yang, "Intelligent fault diagnosis method for rotating machinery via dictionary learning and sparse representation-based classification," *Measurement*, vol. 118, pp. 181–193, Mar. 2018.

[68] S. Wang and H. Chen, "A novel deep learning method for the classification of power quality disturbances using deep convolutional neural network," *Appl. Energy*, vol. 235, pp. 1126–1140, Feb. 2019.

[69] X. Xu, H. Zheng, Z. Guo, X. Wu, and Z. Zheng, "SDD-CNN: Small data-driven convolution neural networks for subtle roller defect inspection," *Appl. Sci.*, vol. 9, no. 7, p. 1364, 2019.

[70] D. Li, D. Chen, B. Jin, L. Shi, J. Goh, and S.-K. Ng, "MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2019, pp. 703–716.

[71] 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. 2018.

[72] Z. Wang, J. Wang, and Y. Wang, "An intelligent diagnosis scheme based on generative adversarial learning deep neural networks and its application to planetary gearbox fault pattern recognition," *Neurocomputing*, vol. 310, pp. 213–222, Oct. 2018.

[73] M. Xia, X. Zheng, M. Imran, and M. Shoaib, "Data-driven prognosis method using hybrid deep recurrent neural network," *Appl. Soft Comput.*, vol. 93, Aug. 2020, Art. no. 106351.

[74] C. K. M. Lee, K. K. H. Ng, C.-H. Chen, H. C. W. Lau, S. Y. Chung, and T. Tsoi, "American sign language recognition and training method with recurrent neural network," *Expert Syst. Appl.*, vol. 167, Apr. 2021, Art. no. 114403.

[75] R. Zhu, X. Tu, and J. X. Huang, "Deep learning on information retrieval and its applications," in *Deep Learning for Data Analytics*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 125–153.

[76] X. Li, X. Ma, F. Xiao, F. Wang, and S. Zhang, "Application of gated recurrent unit (GRU) neural network for smart batch production prediction," *Energies*, vol. 13, no. 22, p. 6121, Nov. 2020.

[77] V. Veerasamy, N. I. A. Wahab, M. L. Othman, S. Padmanaban, K. Sekar, R. Ramachandran, H. Hizam, A. Vinayagam, and M. Z. Islam, "LSTM recurrent neural network classifier for high impedance fault detection in solar PV integrated power system," *IEEE Access*, vol. 9, pp. 32672–32687, 2021.

[78] L. Guo, N. Li, F. Jia, Y. Lei, and J. Lin, "A recurrent neural network based health indicator for remaining useful life prediction of bearings," *Neurocomputing*, vol. 240, pp. 98–109, May 2017.

[79] M. Xia, H. Shao, X. Ma, and C. W. de Silva, "A stacked GRU-RNN-Based approach for predicting renewable energy and electricity load for smart grid operation," *IEEE Trans. Ind. Informat.*, vol. 17, no. 10, pp. 7050–7059, Oct. 2021.

[80] T. Nayak and H. T. Ng, "Effective modeling of encoder-decoder architecture for joint entity and relation extraction," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, no. 5, pp. 8528–8535.

[81] S. Seo, C. Kim, H. Kim, K. Mo, and P. Kang, "Comparative study of deep learning-based sentiment classification," *IEEE Access*, vol. 8, pp. 6861–6875, 2020.

[82] P. Malhotra, V. Tv, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, "Multi-sensor prognostics using an unsupervised health index based on LSTM encoder-decoder," 2016, *arXiv:1608.06154*. [Online]. Available: http://arxiv.org/abs/1608.06154

[83] H. Zhao, S. Sun, and B. Jin, "Sequential fault diagnosis based on LSTM neural network," *IEEE Access*, vol. 6, pp. 12929–12939, 2018.

[84] A. Y. Appiah, X. Zhang, B. B. K. Ayawli, and F. Kyeremeh, "Long short-term memory networks based automatic feature extraction for photovoltaic array fault diagnosis," *IEEE Access*, vol. 7, pp. 30089–30101, 2019.

[85] J. Chung, S. Ahn, and Y. Bengio, "Hierarchical multiscale recurrent neural networks," 2016, *arXiv:1609.01704*. [Online]. Available: http://arxiv.org/abs/1609.01704

[86] A. Dairi, F. Harrou, Y. Sun, and S. Khadraoui, "Short-term forecasting of photovoltaic solar power production using variational auto-encoder driven deep learning approach," *Appl. Sci.*, vol. 10, no. 23, p. 8400, Nov. 2020.

[87] J. P. C. Chiu and E. Nichols, "Named entity recognition with bidirectional LSTM-CNNs," *Trans. Assoc. Comput. Linguistics*, vol. 4, pp. 357–370, Dec. 2016.

[88] P. Park, P. D. Marco, H. Shin, and J. Bang, "Fault detection and diagnosis using combined autoencoder and long short-term memory network," *Sensors*, vol. 19, no. 21, p. 4612, Oct. 2019.

[89] Z. Yang, D. Gjorgjevikj, J. Long, Y. Zi, S. Zhang, and C. Li, "Sparse autoencoder-based multi-head deep neural networks for machinery fault diagnostics with detection of novelties," *Chin. J. Mech. Eng.*, vol. 34, no. 1, pp. 1–12, Dec. 2021.

[90] R. Li, Q. Wu, J. Liu, Q. Wu, C. Li, and Q. Zhao, "Monitoring depth of anesthesia based on hybrid features and recurrent neural network," *Frontiers Neurosci.*, vol. 14, p. 26, Feb. 2020.

[91] C.-H. Chang, "Deep and shallow architecture of multilayer neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2477–2486, Oct. 2015.

[92] L. Wang, H. Qiu, P. Yang, and L. Mu, "Arc fault detection algorithm based on variational mode decomposition and improved multi-scale fuzzy entropy," *Energies*, vol. 14, no. 14, p. 4137, Jul. 2021.

[93] R. Thirukovalluru, S. Dixit, R. K. Sevakula, N. K. Verma, and A. Salour, "Generating feature sets for fault diagnosis using denoising stacked auto-encoder," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2016, pp. 1–7.

[94] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. 25th Int. Conf. Mach. Learn. (ICML)*, 2008, pp. 1096–1103.

[95] N. Tavakoli, S. Siami-Namini, M. A. Khanghah, F. M. Soltani, and A. S. Namin, "An autoencoder-based deep learning approach for clustering time series data," *Social Netw. Appl. Sci.*, vol. 2, no. 5, pp. 1–25, May 2020.

[96] J. Wu, Z. Yan, and Q. Sun, "Multiple faults detection of three-level NPC inverter based on improved deep learning network," in *Proc. Int. Conf. Appl. Techn. Cyber Secur. Intell.* Cham, Switzerland: Springer, 2019, pp. 1575–1583.

[97] M.-Y. Chow, P. Mangum, and R. J. Thomas, "Incipient fault detection in DC machines using a neural network," in *Proc. 22nd Asilomar Conf. Signals, Syst. Comput.*, vol. 2, 1988, pp. 706–709.

[98] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, no. 12, pp. 1–38, 2010.

[99] S. Nagpal, M. Singh, R. Singh, and M. Vatsa, "Regularized deep learning for face recognition with weight variations," *IEEE Access*, vol. 3, pp. 3010–3018, 2015.

[100] J. F. Bermejo, J. F. G. Fernández, F. O. Polo, and A. C. Márquez, "A review of the use of artificial neural network models for energy and reliability prediction. A study of the solar PV, hydraulic and wind energy sources," *Appl. Sci.*, vol. 9, no. 9, p. 1844, May 2019.

[101] D. Qian and W. K. Cheung, "Enhancing variational autoencoders with mutual information neural estimation for text generation," in *Proc. Conf. Empirical Methods Natural Lang. Process., 9th Int. Joint Conf. Natural Lang. Process. (EMNLP-IJCNLP)*, 2019, pp. 4038–4048.

[102] X. Li, X.-D. Jia, W. Zhang, H. Ma, Z. Luo, and X. Li, "Intelligent cross-machine fault diagnosis approach with deep auto-encoder and domain adaptation," *Neurocomputing*, vol. 383, pp. 235–247, Mar. 2020.

[103] M. Ma, C. Sun, and X. Chen, "Deep coupling autoencoder for fault diagnosis with multimodal sensory data," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 1137–1145, Mar. 2018.

[104] F. Scheller, B. Burgenmeister, H. Kondziella, S. Kühne, D. G. Reichelt, and T. Bruckner, "Towards integrated multi-modal municipal energy systems: An actor-oriented optimization approach," *Appl. Energy*, vol. 228, pp. 2009–2023, Oct. 2018.

[105] L. Wen, L. Gao, and X. Li, "A new deep transfer learning based on sparse auto-encoder for fault diagnosis," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 1, pp. 136–144, Jan. 2017.

[106] Y. Qi, C. Shen, D. Wang, J. Shi, X. Jiang, and Z. Zhu, "Stacked sparse autoencoder-based deep network for fault diagnosis of rotating machinery," *IEEE Access*, vol. 5, pp. 15066–15079, 2017.

[107] H. Han, L. Hao, D. Cheng, and H. Xu, "GAN-SAE based fault diagnosis method for electrically driven feed pumps," *PLoS ONE*, vol. 15, no. 10, Oct. 2020, Art. no. e0239070.

[108] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mech. Syst. Signal Process.*, vol. 115, pp. 213–237, Jan. 2019.

[109] R. Souriau, J. Lerbet, H. Chen, and V. Vigneron, "A review on generative Boltzmann networks applied to dynamic systems," *Mech. Syst. Signal Process.*, vol. 147, Jan. 2021, Art. no. 107072.

[110] G. E. Hinton and T. J. Sejnowski, "Learning and relearning in Boltzmann machines," *Parallel Distrib. Process., Explor. Microstructure Cognition*, vol. 1, nos. 282–317, p. 2, 1986.

[111] C. Li, R.-V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, and R. E. Vásquez, "Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis," *Neurocomput.*, vol. 168, pp. 119–127, Nov. 2015.

[112] N. Srivastava and R. Salakhutdinov, "Multimodal learning with deep Boltzmann machines," in *Proc. NIPS*, vol. 1, 2012, p. 2.

[113] H. Shao, H. Jiang, H. Zhang, and T. Liang, "Electric locomotive bearing fault diagnosis using a novel convolutional deep belief network," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2727–2736, Mar. 2017.

[114] N. L. Roux and Y. Bengio, "Representational power of restricted Boltzmann machines and deep belief networks," *Neural Comput.*, vol. 20, no. 6, pp. 1631–1649, Jun. 2008.

[115] J. Tao, Y. Liu, and D. Yang, "Bearing fault diagnosis based on deep belief network and multisensor information fusion," *Shock Vib.*, vol. 2016, pp. 1–9, Aug. 2016.

[116] L. Liao, W. Jin, and R. Pavel, "Enhanced restricted Boltzmann machine with prognosability regularization for prognostics and health assessment," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7076–7083, Nov. 2016.

[117] V. T. Tran, F. AlThobiani, and A. Ball, "An approach to fault diagnosis of reciprocating compressor valves using Teager–Kaiser energy operator and deep belief networks," *Expert Syst. Appl.*, vol. 41, no. 9, pp. 4113–4122, Jul. 2014.

[118] J. Zhang, H. Wang, J. Chu, S. Huang, T. Li, and Q. Zhao, "Improved Gaussian–Bernoulli restricted Boltzmann machine for learning discriminative representations," *Knowl.-Based Syst.*, vol. 185, Dec. 2019, Art. no. 104911.

[119] X. Huang, X. Zhang, Y. Xiong, H. Liu, and Y. Zhang, "A novel intelligent fault diagnosis approach for early cracks of turbine blades via improved deep belief network using three-dimensional blade tip clearance," *IEEE Access*, vol. 9, pp. 13039–13051, 2021.

[120] F. Lv, C. Wen, Z. Bao, and M. Liu, "Fault diagnosis based on deep learning," in *Proc. Amer. Control Conf. (ACC)*, Jul. 2016, pp. 6851–6856.

[121] S. Zhao, J. Li, and X. Cheng, "Sensor fault diagnosis based on adaptive arc fuzzy DBN-Petri net," *IEEE Access*, vol. 9, pp. 20305–20317, 2021.

[122] P. B. Dash, B. Naik, J. Nayak, and S. Vimal, "Deep belief network-based probabilistic generative model for detection of robotic manipulator failure execution," *Soft Comput.*, pp. 1–13, Jan. 2021, doi: 10.1007/s00500-021-05572-0.

[123] Y. Zhang, J. Ji, and B. Ma, "Fault diagnosis of reciprocating compressor using a novel ensemble empirical mode decomposition-convolutional deep belief network," *Measurement*, vol. 156, May 2020, Art. no. 107619.

[124] A. Malik, A. Haque, K. S. Bharath, and Z. A. Jaffery, "Transfer learning-based novel fault classification technique for grid-connected PV inverter," in *Innovations in Electrical and Electronic Engineering*. Singapore: Springer, 2021, pp. 217–224.

[125] J. Liu, Q. Zhang, X. Li, G. Li, Z. Liu, Y. Xie, K. Li, and B. Liu, "Transfer learning-based strategies for fault diagnosis in building energy systems," *Energy Buildings*, vol. 250, Nov. 2021, Art. no. 111256.

[126] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, vol. 25, Dec. 2012, pp. 1097–1105.

[127] D. Xiao, Y. Huang, L. Zhao, C. Qin, H. Shi, and C. Liu, "Domain adaptive motor fault diagnosis using deep transfer learning," *IEEE Access*, vol. 7, pp. 80937–80949, 2019.

[128] S. H. Cho, S. Kim, and J.-H. Choi, "Transfer learning-based fault diagnosis under data deficiency," *Appl. Sci.*, vol. 10, no. 21, p. 7768, Nov. 2020.

[129] S. Shao, S. McAleer, R. Yan, and P. Baldi, "Highly accurate machine fault diagnosis using deep transfer learning," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2446–2455, Apr. 2018.

[130] J. Li, Y. Wang, Y. Zi, and S. Jiang, "A local weighted multi-instance multilabel network for fault diagnosis of rolling bearings using encoder signal," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 10, pp. 8580–8589, Oct. 2020.

[131] J. Han, S. H. Miao, H. R. Yin, S. Y. Guo, Z. X. Wang, F. X. Yao, and Y. J. Lin, "Deep-adversarial-transfer learning based fault classification of power lines in smart grid," *IOP Conf. Ser., Earth Environ. Sci.*, vol. 701, no. 1, Mar. 2021, Art. no. 012074.

[132] P. Wang and R. X. Gao, "Transfer learning for enhanced machine fault diagnosis in manufacturing," *CIRP Ann.*, vol. 69, no. 1, pp. 413–416, 2020.

[133] R. Zhang, H. Tao, L. Wu, and Y. Guan, "Transfer learning with neural networks for bearing fault diagnosis in changing working conditions," *IEEE Access*, vol. 5, pp. 14347–14357, 2017.

[134] W. Qian, S. Li, and J. Wang, "A new transfer learning method and its application on rotating machine fault diagnosis under variant working conditions," *IEEE Access*, vol. 6, pp. 69907–69917, 2018.

[135] B. Yang, Y. Lei, F. Jia, N. Li, and Z. Du, "A polynomial kernel induced distance metric to improve deep transfer learning for fault diagnosis of machines," *IEEE Trans. Ind. Electron.*, vol. 67, no. 11, pp. 9747–9757, Nov. 2020.

[136] Z. Wan, R. Yang, M. Huang, N. Zeng, and X. Liu, "A review on transfer learning in EEG signal analysis," *Neurocomputing*, vol. 421, pp. 1–14, Jan. 2021.

[137] D. Zhang and T. Zhou, "Deep convolutional neural network using transfer learning for fault diagnosis," *IEEE Access*, vol. 9, pp. 43889–43897, 2021.

[138] Q. Li, C. Shen, L. Chen, and Z. Zhu, "Knowledge mapping-based adversarial domain adaptation: A novel fault diagnosis method with high generalizability under variable working conditions," *Mech. Syst. Signal Process.*, vol. 147, Jan. 2021, Art. no. 107095.

[139] J. Zhu, N. Chen, and C. Shen, "A new deep transfer learning method for bearing fault diagnosis under different working conditions," *IEEE Sensors J.*, vol. 20, no. 15, pp. 8394–8402, Aug. 2020.

[140] L. Zhang, L. Guo, H. Gao, D. Dong, G. Fu, and X. Hong, "Instance-based ensemble deep transfer learning network: A new intelligent degradation recognition method and its application on ball screw," *Mech. Syst. Signal Process.*, vol. 140, Jun. 2020, Art. no. 106681.

[141] X. Fang, G. Gong, G. Li, L. Chun, W. Li, and P. Peng, "A hybrid deep transfer learning strategy for short term cross-building energy prediction," *Energy*, vol. 215, Jan. 2021, Art. no. 119208.

[142] N. Li, H. Hao, Q. Gu, D. Wang, and X. Hu, "A transfer learning method for automatic identification of sandstone microscopic images," *Comput. Geosci.*, vol. 103, pp. 111–121, Jun. 2017.

[143] Y. Xu, S. J. Pan, H. Xiong, Q. Wu, R. Luo, H. Min, and H. Song, "A unified framework for metric transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 6, pp. 1158–1171, Jun. 2017.

[144] X. Liu, Z. Liu, G. Wang, Z. Cai, and H. Zhang, "Ensemble transfer learning algorithm," *IEEE Access*, vol. 6, pp. 2389–2396, 2017.

[145] H. Chang, J. Han, C. Zhong, A. M. Snijders, and J.-H. Mao, "Unsupervised transfer learning via multi-scale convolutional sparse coding for biomedical applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 5, pp. 1182–1194, May 2017.

[146] Y. Yang, J. Guo, Q. Ye, Y. Xia, P. Yang, A. Ullah, and K. Muhammad, "A weighted multi-feature transfer learning framework for intelligent medical decision making," *Appl. Soft Comput.*, vol. 105, Jul. 2021, Art. no. 107242.

[147] W. Li, S. Gu, X. Zhang, and T. Chen, "Transfer learning for process fault diagnosis: Knowledge transfer from simulation to physical processes," *Comput. Chem. Eng.*, vol. 139, Aug. 2020, Art. no. 106904.

[148] K. Zhang, N. Robinson, S.-W. Lee, and C. Guan, "Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network," *Neural Netw.*, vol. 136, pp. 1–10, Apr. 2021.

[149] S. Chen, H. Ge, H. Li, Y. Sun, and X. Qian, "Hierarchical deep convolution neural networks based on transfer learning for transformer rectifier unit fault diagnosis," *Measurement*, vol. 167, Jan. 2021, Art. no. 108257.

[150] Y. Deng, D. Huang, S. Du, G. Li, C. Zhao, and J. Lv, "A double-layer attention based adversarial network for partial transfer learning in machinery fault diagnosis," *Comput. Ind.*, vol. 127, May 2021, Art. no. 103399.

[151] M. B. H. Moran, M. D. B. Faria, G. A. Giraldi, L. F. Bastos, and A. Conci, "Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs," *Comput. Biol. Med.*, vol. 129, Feb. 2021, Art. no. 104139.

[152] E. Soleimani and E. Nazerfard, "Cross-subject transfer learning in human activity recognition systems using generative adversarial networks," *Neurocomputing*, vol. 426, pp. 26–34, Feb. 2021.

[153] Z. Zhang, X. Li, L. Wen, L. Gao, and Y. Gao, "Fault diagnosis using unsupervised transfer learning based on adversarial network," in *Proc. IEEE 15th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2019, pp. 305–310.

[154] G. Sun, L. Liang, T. Chen, F. Xiao, and F. Lang, "Network traffic classification based on transfer learning," *Comput. Elect. Eng.*, vol. 69, pp. 920–927, Jul. 2018.

[155] Y. Wang, J. Zhai, Y. Li, K. Chen, and H. Xue, "Transfer learning with partial related 'instance-feature' knowledge," *Neurocomputing*, vol. 310, pp. 115–124, Oct. 2018.

[156] S. Lu, Q. He, and J. Zhao, "Bearing fault diagnosis of a permanent magnet synchronous motor via a fast and online order analysis method in an embedded system," *Mech. Syst. Signal Process.*, vol. 113, pp. 36–49, Dec. 2018.

[157] C. Mazo, J. Bernal, M. Trujillo, and E. Alegre, "Transfer learning for classification of cardiovascular tissues in histological images," *Comput. Methods Programs Biomed.*, vol. 165, pp. 69–76, Oct. 2018.

[158] B. Liu, Y. Xiao, and Z. Hao, "A selective multiple instance transfer learning method for text categorization problems," *Knowl.-Based Syst.*, vol. 141, pp. 178–187, Feb. 2018.

[159] M. Ribeiro, K. Grolinger, H. F. ElYamany, W. A. Higashino, and M. A. M. Capretz, "Transfer learning with seasonal and trend adjustment for cross-building energy forecasting," *Energy Buildings*, vol. 165, pp. 352–363, Apr. 2018.

[160] W. Lu, B. Liang, Y. Cheng, D. Meng, J. Yang, and T. Zhang, "Deep model based domain adaptation for fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2296–2305, Mar. 2016.

[161] M. W. Akram, G. Li, Y. Jin, X. Chen, C. Zhu, and A. Ahmad, "Automatic detection of photovoltaic module defects in infrared images with isolated and develop-model transfer deep learning," *Sol. Energy*, vol. 198, pp. 175–186, Mar. 2020.

[162] C. Chen, Z. Li, J. Yang, and B. Liang, "A cross domain feature extraction method based on transfer component analysis for rolling bearing fault diagnosis," in *Proc. 29th Chin. Control Decis. Conf. (CCDC)*, May 2017, pp. 5622–5626.

[163] M. Z. Sheriff, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Fault detection using multiscale PCA-based moving window GLRT," *J. Process Control*, vol. 54, pp. 47–64, Jun. 2017.

[164] R. Fazai, M. Mansouri, K. Abodayeh, V. Puig, M.-I.-N. Raouf, H. Nounou, and M. Nounou, "Multiscale Gaussian process regression-based generalized likelihood ratio test for fault detection in water distribution networks," *Eng. Appl. Artif. Intell.*, vol. 85, pp. 474–491, Oct. 2019.

[165] C. Botre, M. Mansouri, M. N. Karim, H. Nounou, and M. Nounou, "Multiscale PLS-based GLRT for fault detection of chemical processes," *J. Loss Prevention Process Ind.*, vol. 46, pp. 143–153, 2017.

[166] H.-H. Bock and E. Diday, *Analysis of Symbolic Data: Exploratory Methods for Extracting Statistical Information From Complex Data*. Springer, 2012.

[167] M.-F. Harkat, M. Mansouri, M. N. Nounou, and H. N. Nounou, "Fault detection of uncertain chemical processes using interval partial least squares-based generalized likelihood ratio test," *Inf. Sci.*, vol. 490, pp. 265–284, Jul. 2019.

[168] A. Kouadri, M. Hajji, M.-F. Harkat, K. Abodayeh, M. Mansouri, H. Nounou, and M. Nounou, "Hidden Markov model based principal component analysis for intelligent fault diagnosis of wind energy converter systems," *Renew. Energy*, vol. 150, pp. 598–606, May 2020.

MAJDI MANSOURI (Senior Member, IEEE) received the degree in electrical engineering from SUPCOM, Tunis, Tunisia, in 2006, the M.Sc. degree in electrical engineering from ENSEIRB, Bordeaux, France, in 2008, the Ph.D. degree in electrical engineering from University of Technology of Troyes, Troyes, France, in 2011, and the H.D.R. (Accreditation To Supervise Research) degree in electrical engineering from the University of Orléans, France, in 2019. In 2011, he joined the Electrical and Computer Engineering Program, Texas A&M University at Qatar, where he is currently an Associate Research Scientist. He is the author of more than 150 publications. He is also the author of the book *Data-Driven and Model-Based Methods for Fault Detection and Diagnosis* (Elsevier, 2020). His research interests include development of model-based, data-driven, and machine learning techniques for fault detection and diagnosis.

MOHAMED TRABELSI (Senior Member, IEEE) received the B.Sc. degree in electrical engineering from INSAT, Tunisia, in 2006, and the M.Sc. degree in automated systems and the Ph.D. degree in energy systems from INSA Lyon, France, in 2006 and 2009, respectively. From October 2009 to August 2018, he held different research positions at Qatar University and Texas A&M University at Qatar. Since September 2018, he has been an Associate Professor with Kuwait College of Science and Technology. He has published more than 100 journals and conference papers. He is the author of two books and two book chapters. His research interests include systems control with applications arising in the contexts of power electronics, energy conversion, renewable energies integration, and smart grids.
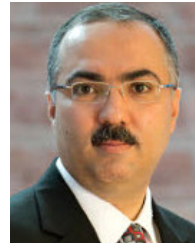
MOHAMED NOUNOU (Senior Member, IEEE) is currently a Professor of chemical engineering at TAMU-Texas A&M University at Qatar. He has more than 19 years of combined academic and industrial experience. He has successfully served as the lead PI and a PI on several QNRF projects (six NPRP projects and three UREP projects). He has published more than 200 refereed journals and conference publications and book chapters. His research interests include systems engineering and control, with emphasis on process modeling, monitoring, and estimation. He is a Senior Member of the American Institute of Chemical Engineers (AIChE).

HAZEM NOUNOU (Senior Member, IEEE) is currently a Professor of electrical and computer engineering, Texas A&M University at Qatar. He has more than 19 years of academic and industrial experience. He has significant experience in research on control systems, data-based control, system identification and estimation, fault detection, and system biology. He has been awarded several NPRP research projects in these areas. He has successfully served as the lead PI and a PI on five QNRF projects, some of which were in collaboration with other PIs in this proposal. He has published more than 200 refereed journals and conference papers and book chapters. He has served as an Associate Editor and on the technical committees for several international journals and conferences.

· · ·