Renewable and Sustainable Energy Reviews 138 (2021) 110512

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: http://www.elsevier.com/locate/rser

Application of Artificial Neural Networks to photovoltaic fault detection and diagnosis: A review

B. Li [a,b], C. Delpha [b], D. Diallo [a,c,*], A. Migan-Dubois [a]

[a] Université Paris-Saclay, CentraleSupélec, CNRS, Sorbonne Université, GeePs, Gif Sur Yvette, 91192, France
[b] Université Paris-Saclay, CNRS, CentraleSupélec, L2S, Gif Sur Yvette, 91192, France
[c] Shanghai Maritime University, Shanghai, 201306, China

ARTICLE INFO

Keywords:
Photovoltaic
Artificial neural network
Fault detection
Fault classification
Machine learning
Deep learning

ABSTRACT

The rapid development of photovoltaic (PV) technology and the growing number and size of PV power plants require increasingly efficient and intelligent health monitoring strategies to ensure reliable operation and high energy availability. Among the various techniques, Artificial Neural Network (ANN) has exhibited the functional capacity to perform the identification and classification of PV faults. In the present review, a systematic study on the application of ANN and hybridized ANN models for PV fault detection and diagnosis (FDD) is conducted. For each application, the targeted PV faults, the detectable faults, the type and amount of data used, the model configuration and the FDD performance are extracted, and analyzed. The main trends, challenges and prospects for the application of ANN for PV FDD are extracted and presented.

1. Introduction

In recent decades, photovoltaic (PV) technology has experienced an accelerated development as one of the promising renewable energy sources relying on various merits, including pollution-free, safe energy generation, noiseless operation and decreasing installation costs [1–3]. Nevertheless, due to the outdoor operating conditions (random variations of environmental conditions) and the potential damages involved in the manufacturing, transportation or installation [4], various PV faults may accordingly arise up and lead to different levels of degradation, power loss or even fire hazard [5,6]. It is reported that the power loss in the UK could reach 18.9% in one year [7], which is found mainly due to the sustained inverter failure and shading fault. As a consequence, it is relevant to conduct efficient health monitoring for PV power plants to ensure the reliability and durability of energy production.

Typical PV fault detection and diagnosis (FDD) strategies can be broadly classified into two categories, visual inspection and automatic fault analysis. The automatic ones could be realized via various methodologies, where the most popular ones include model-based residual analysis and data-driven methods. For the latter one, several kinds of data are commonly used, like electrical measurements, environmental data or images of PV panels. These data-driven analyses can be done with different techniques, such as statistical methods or machine learning technology (MLT) [8]. Compared to others, MLT is competent and proficient to deal with complex and non-linear problems [9].

MLTs applied for FDD consist of various methods with distinctive principles and structures. The most common ones include Artificial Neural Network (ANN) [10], Fuzzy Logic (FL) [11], Support Vector Machine (SVM) [12], *k*-Nearest Neighbor algorithm (*k*NN) [13] and Decision Tree (DT) [14]-based techniques (including random forest (RF) [15]). Through keyword research[1] and the subsequent content verification in common publishers or research platforms (e.g., Science Direct, IEEE Xplore, Google Scholar, Research gate), the number of reported publications on PV FDD from 2009 to July 2020, for different types of MLTs are summarized and presented in Fig. 1.

It can be observed that ANN technique has drawn more research interest during the last decade. All the applications differ in various aspects e.g. input data, data pre-processing, ANN type and structure (*Shallow Neural Network* (SNN) and *Deep Neural Network* (DNN) [16],[2]),

* Corresponding author. Université Paris-Saclay, CentraleSupelec, CNRS, GeePs, Sorbonne Université, 3 rue Joliot Curie, 91192 Gif Sur Yvette, France.
E-mail address: demba.diallo@geeps.centralesupelec.fr (D. Diallo).

[1] Setting of keywords for different MLTs: 'Photovoltaic' +'fault' or 'defect' or 'detection' or 'diagnosis' +technique name ('neural network'/'fuzzy logic'/'support vector machine'/'kNN'/'decision tree'/'random forest').

[2] SNN and DNN are the 2 categories of ANN, which are distinguished according to the model's depth, i.e., the number of hidden layers. However, it should be noted that this number of layers is a relative notion that principally depends on the models to be compared [16]. Hereafter, based on an overall analysis, we set 3 hidden layers as the bound, so as to facilitate the presentation and performance comparison presented in the next section.

https://doi.org/10.1016/j.rser.2020.110512
Received 27 December 2019; Received in revised form 21 October 2020; Accepted 22 October 2020
Available online 6 November 2020
1364-0321/© 2020 Elsevier Ltd. All rights reserved.

Nomenclature

Techniques related terminology

1D, 2D	1 Dimension, 2 Dimension
ANN	Artificial Neural Network
ART	Adaptive Resonance Theory
BP	Back Propagation
CNN	Convolutional Neural Network
DBN	Deep Belief Network
DnCNN	Denoising CNN
DNN	Deep Neural Network
DT	Decision Tree
DWT	Discrete Wavelet Transform
ENN	Extension Neural Network
FDD	Fault Detection and Diagnosis
FL	Fuzzy Logic
GA	Genetic Algorithm
GAN	Generative Adversarial Network
GK-FCM	Gaussian Kernel function-based Fuzzy C-Means
GLCM	Grey Level Co-occurrence Matrix
I/O	Input and Output
KELM/ELM	(Kernel based) Extreme Learning Machine
*k*NN	*k*-Nearest Neighbors
LAPART	Laterally Primed Adaptive Resonance Theory
LDA	Linear Discriminant Analysis
LM	Levenberg-Marquardt
LRMR	Low Rank Matrix Recovery
MSE	Mean Squared Error
MLP	Multi-layer Perception
MLT	Machine Learning Technique
MRA	Multi Resolution Analysis
MSDP	Multistate Data Processing
NAG	Nesterov Accelerated Gradient
NMS	Nelder-Mead Simplex
NN	Neural Network
PCA	Principal Component Analysis
PSO	Particle Swarm Optimization
PNN	Probabilistic Neural Network
ReLU	Rectified Linear Unit
ResNet	Deep Residual Network
RBF	Radical Basis Function
RBM	Restricted Boltzmann Machine
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
SA	Simulated Annealing
SCG	Scaled Conjugate Gradient
SGD	Stochastic Gradient Descent
SNN	Shallow Neural Network
SVM	Support Vector Machine
t-SNE	*t*-Distributed Stochastic Neighbor Embedding
WNN	Wavelet neural network

Photovoltaic related terminology

AC	Alternating Current
AF	Arc Fault
BPD	Bypass diode
DC	Direct Current
EL	Electroluminescence
FF	Fill factor
GF	Ground Fault
IGBT	Insulated Gate Bipolar Transistor
IR	Infrared thermography
LLF	Line-to-line fault
mc-Si	Multicrystalline Silicon
MPP(T)	Maximum Power Point (Tracking)
OC	Open Circuit
PID	Potential Induced Degradation
PS	Partial Shading
PV	Photovoltaic
SC	Short Circuit
sc-Si	Single crystalline Silicon
UAV	Unmanned Aerial Vehicle

Symbol

C_I	Classification coefficient of current
C_V	Classification coefficient of voltage
C_R	Resistance coefficient
Ep	RMSE of parameter identification
fval	Optimal fitness value of I–V curve
G	Solar irradiance (W/m^2)
I	Output current (A)
I_{AC}	AC side current (A)
I_{load}	Load current (V)
I_{MPP}	Current at MPP (A)
I_{ph}	Photocurrent (A)
I_{SC}	Short circuit current (A)
n	Ideality factor
Np	Number of ANN parameters
P_{MPP}	Power at MPP (W)
PR	Power ratio (%)
$R_{I\text{-}MPP}$	Ratio of current at MPP (%)
R_S	Series resistance (Ω)
$R_{V\text{-}MPP}$	Ratio of voltage at MPP
R_{Voc}	Ratio of open circuit voltage (%)
T_A	Ambient temperature (°C)
T_M	Module temperature (°C)
V	Output voltage (V)
V_{AC}	Inverter voltage (V)
V_{MPP}	AC side voltage (V)
V_{load}	Load voltage (V)
V_{OC}	Open circuit voltage (V)
VR	Voltage ratio (%)
v_{WIND}	Wind speed (m/s)
d_{WIND}	Wind direction

parameter configuration, hybrid application and performance. Despite this abundant literature, there is no analysis and comparison of the applications of ANN for PV FDD. This paper is an attempt to fill this gap by doing a review of the papers published from 2009 to July 2020.

As a hierarchical model, one neural network (NN) generally includes one input layer, several hidden layers, and one output layer. Each layer is composed of several connected units (named neurons), each one associated with an activation function. It operates as parallelized processors to deal with complex systems [17]. ANN has various merits, like an excellent approximation of nonlinear function, fast decision making, no restriction on the normality or independence of input [17,18]. Thus, ANN has also been widely exploited in other PV domains [19], including solar radiation prediction [20], MPP tracking [21], solar energy system modeling [18], PV system sizing [22], and performance prediction of solar collector system [23].

This paper is interested in carrying out a literature review to answer the following questions regarding the use of ANN for PV FDD:

- How ANN techniques are integrated in the PV FDD?
- What types of PV faults are detectable using ANN-based techniques?

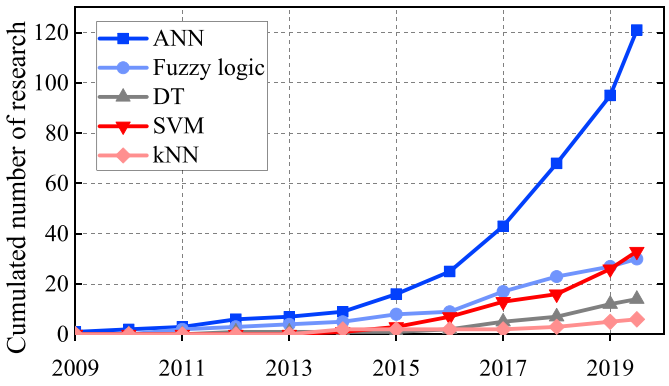


Fig. 1. Evolution of reported applications of MLTs for PV FDD.

- What are the performance of these techniques when applied for PV FDD?
- What are the challenges and prospects that ANN would face for PV health monitoring?

The contribution of this paper is reflected in:

1) Applications of ANN for PV FDD are, to our best knowledge, firstly summarized in detail based on the type of neural networks: shallow, deep, and hybridized. For each type of application, their pros and limitations have been identified.
2) The common trends for the selection and the configuration of one ANN model when using 1-dimension (1D) or 2-dimension (2D) PV data or when targeting at different types of PV faults have been highlighted.
3) Four reported public PV databases, including both 1D data or 2D images, are identified from the literature and summarized.
4) Common challenges (e.g. the model configuration, availability of public database) and common prospects (e.g. the candidate models, feature transformation between1D and 2D, real-time health monitoring) have also been highlighted.

The remainder of this paper is structured as follows: Section 2 presents a systematic review of the application of SNNs, DNNs and hybridized methods for PV FDD, where the key issues like model configuration, targeted fault, performance are all detailed and analyzed. Then, based on the aforementioned four questions, section 3 discusses the reviewed cases, and some common challenges and prospects are also highlighted. Section 4 concludes the paper.

2. Application of ANN for PV FDD

This section presents an insight into the various applications of ANN for PV FDD by separating the adopted models into three main groups, SNN, DNN and hybridized neural network (NN).

2.1. Application of SNN for PV FDD

Two categories of SNN applications are considered in this subsection, i.e., direct FDD (end-to-end type, where the SNNs output straightforward the fault type) and indirect FDD (where the output data of SNN models need further interpretation to identify the fault type).

2.1.1. Direct FDD

In the literature, various types of SNN models have been applied for direct FDD such as *Multilayer Perceptron neural network* (MLP) [24], *Radial basis function neural network* (RBF) [25], *Probabilistic neural network* (PNN) [26], *Extension neural network* (ENN) [27], *Extreme learning machine* (ELM) [28], *Elman neural network* (Elman NN) [29], and *Wavelet neural network* (WNN) [30].

In most cases, different factors (like the network structure, the number of parameters *Np* (neuron weights and bias), input/output (I/O) setting, activation function, tuning algorithm, PV technology and scale) have been taken into account to reach the best-expected performance. These applications, sorted by model type, are presented in Table 1 followed by a discussion.

Focusing on the four questions proposed in the introduction, the applications displayed in Table 1 are now analyzed:

1) Integration of SNN in PV FDD
   - *ANN type:* Ranking the selected models according to their number of uses, MLP comes first (62.5%, 15 out of 24 application cases) followed by PNN, RBF and ENN.
   - *Model structure:* Relatively simple structures are most often considered, 63.6% (21 out of 33 models) with 1 hidden layer and the remaining with 2 hidden layers. The number of neurons in the hidden layer is generally less or around 10, while in some models more neurons are selected [36,41,42]. As for the number of parameters, its value is generally less than 500 in the reviewed cases.
   - *Amount of data:* Total data volumes used are in the order of hundreds or thousands, and there is usually more data for training than for testing (85.7%), while the dataset for validation is relatively less adopted (4 out of 24 cases). In term of data type, simulation data is the most common one (56.5%, 35 out of 62 reported datasets for training, validation or testing) compared to field-measured data (43.5%). This is because simulation data is easier to obtain despite it may not fully represent actual data due to modeling errors and uncertainties and actual environmental conditions.
   - *I/O setting:* For the input, measured features (like environmental and electrical data) are preferred to calculated ones (like the ratio of voltage or current). V_{MPP}, I_{MPP}, G, T_M are the most common ones. The selection of features depends on the target fault and PV plant scale, i.e., module or array level. For the output, some cases (like in Ref. [32,33]) use the combination binary output values of each neuron to determine the fault type. In some cases, authors export the feature values that correspond directly either to the fault type (like in Refs. [38,50]) or to the fault severity (like in Ref. [35,42]), while the others do not specify.
   - *Activation function:* Only in 11 out of 24 cases, the activation function used is specified. In the hidden layer, non-linear functions (like *tanh* and *logsig*) are commonly adopted, which permits the learning of complex relationship. As for the output layer, linear functions (like *pureline*, softmax) are preferred, which allows the multiple outputs for fault classification.
   - *PV technology and scale:* Almost all PV devices tested are in crystalline silicon technology, if specified. This corresponds well to its representation in terms of installed capacity worldwide. Regarding the scale of the platform under test, string or array level are the most common types (23 out of 24 cases), while the module level is less investigated. The reported power rating varies from 40Wp to 29 kWp.
2) Detectable faults

Among the targeted PV faults, short circuit (SC), partial shading (PS) and abnormal aging are the most studied ones. The particularity of these faults is their significant impact on the electrical features. Other faults like soiling, line-line fault (LLF), and failures in components on the DC side are relatively less investigated.

3) Performance

Generally, diagnosis accuracy is higher than 90%. However, it should be noted that comparison of performance between different research results is complex due to the difference in the database, model

Table 1
Details of PV FDD research using direct SNN.

Year & Ref	ANN type, Structure & Number of parameters	Amount of data			I/O setting		Activation function	PV techno. & scale	FDD Accuracy	Target Fault	Notes
		Train	Validate	Test	Input	Output					
2009 [31]	MLP 5-9-8 *Np* = 139	NC (*simu*)	–	NC (*simu*)	• 5 failure modes	• 8 causes	•tanh (hidden layer) •logsig (output layer)	NC, PV array	**100%** (average)	• Overheating • Power loss • Driven failure • Abnormal voltage • Abnormal meter	•Input data normalized to [-1,1]
2014 [32]	MLP 3-4-3 *Np* = 34	200 (*exp*)	–	4 (*exp*)	• T_M • V_{MPP} (modu) • I_{MPP} (modu)	•3 faults (by 3 binary codes)	•tanh (hidden layer)	sc-Si, JHX100 M, 1×72 PV string (3.9 kWp)	**90%** (average)	• Cell crack • PS • Aging & SC module	
2015 [33]	MLP 4-11-4 *Np* = 107	1600 (*simu*)	–	400 (*simu*)	• G • T_A • V_{MPP} (array) • I_{MPP} (array)	• Healthy •3 faults (by 4 binary codes)	• NC	NC, PV array	**91.94%** (average) 100% (Healthy) 97.95% (Aging) 78.69% (SC) 100% (PS)	• Aging • PS • SC module	Trained by Levenberg-Marquardt (LM) back propagation (BP) [24] algorithm [70]
2016 [34]	MLP 4-NC-3 *Np* NC	992 (*simu*)	–	NC (*exp*)	• I_{MPP} (array) • V_{MPP} (array) • I_{SC} (array) • V_{OC} (array)	• Healthy • 2 faults	NC	NC, PV array	**30%** (1 SC module) **85%** (2 or more SC modules) **~100%** (Aging, Rs>4Ω)	• SC module • Aging	
2016 [35]	MLP 2-20-1 *Np* = 83	7819 (*simu*)	–	50 (*simu*)	• ΔP (string) • ΔI_{SC} (string)	• PS	NC	NC, 1×16 PV string	Training MSE: **0.058**	• PS (exist or not)	
	MLP 2-NC-1 *Np* NC	12,495 (*simu*)	–	50 (*simu*)	• ΔP (string) • P_{MPP} (string)	• PS factor	NC		Training MSE: **0.067**	• PS (with shading factor)	
	MLP 1-NC-1 *Np* NC	28,512 (*simu*)	–	10 (*simu*)	•PS factor	• Shaded module number	NC		Training MSE: **0.788**	• PS (with number of shaded modules)	
2016 [36]	MLP 3-13-13-1 *Np* = 251 RBF 3-49-4 *Np* = 399	620 (*simu*)	–	155 (*simu*)	• R_{Voc} (array) • $R_{I\text{-}MPP}$ (array) • $R_{V\text{-}MPP}$ (array)	• 4 faults	•*logsig* (hidden layer)	mc-Si, 1×4 PV string (480Wp)	**90.3%** (average) **68.4%** (average)	Cell/diode/module • Shunted • SC • Inversed • Connection resistance fault	• Trained by LM BP algorithm
2017 [37]	MLP 2-8-8-1 *Np* = 107	600 (*simu*)	200 (*simu*)	200 (*simu*)	• G • I_{MPP} (array) • T_M • V_{MPP} (array)	• C_I • C_V	•*logsig* (hidden layer)	sc-Si, Isofoton, 2×8 PV array (1.7 kWp)	NC	• 1 SC module • 4 SC modules • Faulty string	•Fault type co-determined by C_I and C_V • Trained by LM BP algorithm
2017 [38]	MLP 6-7-12-1 *Np* = 164	NC (*exp*)	–	NC (*exp*)	• G • T_M • V_{OC} (modu) • I_{SC} (modu) • V_{MPP} (modu) • I_{MPP} (modu)	• 4 faults	NC	sc-Si, BP-MSX120, 4×18 PV array (8.6Wp)	**79.86%** (average)	Cell/diode/module • Shunted • SC • Inversed • Connection resistance fault	
	MLP 6-7-11-1 *Np* = 155					• 3 faults			**90.79%** (average)	PS & BPD as •OC • Inversed • Shunted	
2017 [39]	MLP 2-10-3 *Np* = 65	70% (*simu*)	–	30% (*simu*)	• V_{MPP} (string) • I_{MPP} (string)	• 5 faults • (by 3 binary codes)	•*logsig* (hidden layer) •*purelin* (output layer)	NC, 1×3 PV string	NC	• 1 or 2 modules inversed PS (3 levels)	• Trained by LM BP algorithm

(*continued on next page*)

Table 1 (*continued*)

Year & Ref	ANN type, Structure & Number of parameters	Amount of data			I/O setting		Activation function	PV techno. & scale	FDD Accuracy	Target Fault	Notes
		Train	Validate	Test	Input	Output					
2018 [40]	MLP 6-25-4 $Np = 285$	1397 (*exp*)	–	394 (*exp*)	• V_{MPP} (string) • I_{MPP} (string) • V_{load} • I_{load} • V_b (battery) • I_b (battery)	• Healthy • 3 faults	•*tanh* (hidden layer) •*logsig* (output layer)	mc-Si, 1×2 PV string (180Wp)	**97.4%** (average)	• SC module • OC module • SC battery	•Input data normalized to [-1,1]
2019 [41]	MLP 7-33-6 $Np = 475$	28,800 (*simu*)	–	4800 (*simu*)	• G • T_A • V_{MPP} (array) • I_{MPP} (string1, string 2) • V_{AC} (inverter) • I_{AC} (inverter)	• Healthy • 5 faults	NC	mc-Si, 2×8 PV array (5.3 kWp)	**99.7%** (average)	• PS • Aging • SC module • OC string • Inverter fault	• Compared and outperforms DT (accu: 89.9%), kNN (82.5%) and SVM (98.6%)
2019 [42]	MLP 6-35-1 $Np = 287$	80% (*exp*)	10% (*exp*)	10% (*exp*)	• G • T_A • v_{WIND} • d_{WIND} • Humidity • Rainfall	• Soiling rate	•tanh (hidden layer) •purelin (output layer)	NC, PV module	**R^2=92.8%** (fitting rate)	• Soiling	• Input data normalized to [0.1, 0.9]
2020 [43]	MLP 5-10-16 $Np = 241$	4368 (*simu*)	936 (*simu*)	936 (*simu*)	• G • T_M • V_{MPP} (array) • I_{MPP} (array) • P_{MPP} (array)	• 16 faults	•tanh (hidden layer) •logsig (output layer)	2 types (sc-Si and thin film), 9×7 PV array (9.1 kWp)	**99.6%** (average)	• SC • OC • Mismatch • Multi faults	• Scaled Conjugate Gradient (SCG) [44] algorithm is adopted
2020 [45]	MLP 2-NC-2 *Np* NC	NC (*exp*)	–	NC (*exp*)	Signal strength indicator of 2 antennas	• Distance between antennas to AF	NC	NC, PV simulator (2.7 kWp)	**Planar location error<0.2m** (within a range of 4 m)	• AF	• Data augmented • Bayesian regularized
2020 [46]	MLP 8-5-1 $Np = 59$	1200 in total (*simu*)			• G • T_M (average) • V_{MPP} (array) • I_{MPP} (array) • P_{MPP} (array) • FF (array) • I_{SC} (array) • V_{OC} (array)	• 6 faults	•tanh (hidden layer)	NC, 6×7 PV array (3.5 kWp)	**99.9%** (average)	• PS (2 types) • SC module • LLF • BPD SC • Temperature non-uniformity	•Input data normalized to [0, 1] • Compared with DT (accu: 95.9%), RF (96.6%)
2018 [47]	RBF 2-5-5-5 $Np = 77$	4536 (*exp*)	648 (*exp*)	1296 (*exp*)	• PR (string) • VR (string)	• 5 faults	NC	mc-Si, SMT6(60), 1×5 PV string (1.1 kWp)	**77.7%** (average)	• Only PS • 1 to 4 faulty modules without PS	• Input data normalized to [-1,1]
	RBF 2-7-7-9 $Np = 151$					• 9 faults			**92.1%** (average)	• Only PS • 1 to 4 faulty modules without PS • 1 to 4 faulty modules with PS	
2020 [48]	RBF 2-10-1 $Np = 43$ MLP 2-10-10-1 $Np = 153$	97,200 in total (*exp*)			• G • P_{MPP} (string)	• Healthy • 9 faults	NC	mc-Si, SMT6(60), 1×10 PV string (2.2 kWp)	**98.6%** (average) **98.9%** (average)	• OC module	•Input data normalized to [0, 1]
		160 (*simu*)	–		• G	• Healthy	NC			• OC (2 types)	

(*continued on next page*)

Table 1 (*continued*)

Year & Ref	ANN type, Structure & Number of parameters	Amount of data			I/O setting		Activation function	PV techno. & scale	FDD Accuracy	Target Fault	Notes
		Train	Validate	Test	Input	Output					
2015 [49]	PNN 4-4-4-1 *Np* = 49			340 (*simu*)	• T_A • V_{MPP} (array) • I_{MPP} (array)	• 3 faults		mc-Si, TSM-290PC14, 5×20 PV array (29 kWp)	**98.53%** (average) 100% (Healthy) 97.6% (OC50) 96.5% (OC75) 100% (LLF)	• LLF	• OC50/75 means 50%/75% or less PV modules are OC
2017 [50]	PNN 4-4-2-1 *Np* = 37	2224 (*simu*) 2224 (*simu*)	–	736 (*simu*) NC (*exp*)	• G • T_M • V_{MPP} (array) • I_{MPP} (array)	• Healthy • Faulty	NC	sc-Si, Isofoton, 2×15 PV array (9.5 kWp)	**100%** **82.3%**	• 3 SC modules • 10 SC modules • 1 OC string	• For fault detection
	PNN 4-4-3-1 *Np* = 43	1668 (*simu*) 1668 (*simu*)		552 (*simu*) NC (*exp*)	• G • T_M • V_{MPP} (array) • I_{MPP} (array)	• 3 faults			**100%** **98.2%**		• For fault diagnosis
2019 [51]	PNN 2-2-5-1 *Np* = 29	400 (*exp*)	–	500 (*exp*)	• V_{MPP} (array) • I_{MPP} (array)	• Healthy • 4 faults	NC	mc-Si, 4×3 PV array (2.9 kWp)	**97%** (average)	• SC module • OC string • LLF • Multi faults	• Input data normalized by STC value
2020 [52]	PNN 7-8-8-3 *Np* = 170	3000 (*exp*)	–	600 (*exp*)	• G • T_M (average) • V_{MPP} (string1) • V_{MPP} (string2) • I_{MPP} (string1) • I_{MPP} (string2) • Weather	• Healthy • 2 faults	•ReLU (hidden layer) •softmax (output layer)	NC, 2×3 PV array (1.8 kWp)	**100%** (average)	• LLF • OC string	• Use Adam as optimizer • *k*-fold [53] validation • Input data normalized
2010 [54]	ENN 4–10 *Np* = 54 MLP 4-9-10 *Np* = 149	1995 (*simu*)	–	1995 (*simu*)	• P_{MPP} (array) • V_{MPP} (array) • I_{MPP} (array) • V_{OC} (array)	• Healthy • 9 faults	• NC	sc-Si, NT-R5E3E, 2×9 PV array (3.2 kWp)	**100%** (average) **93.3%** (average)	• 1 to 3 module faults in 1 or 2 string	
2014 [55]	WNN 4-12-5 *Np* = 129 MLP 4-12-5 *Np* = 129	400 (*exp*)	–	100 (*exp*)	• V_{OC} (array) • I_{SC} (array) • V_{MPP} (array) • I_{MPP} (array)	• Healthy • 4 faults	•*sigmoid* (output layer)	sc-Si, PV array	**96%** (average) **79%** (average)	• SC • OC • PS • Aging	• Use Gaussian function for hidden layer
2018 [56]	Elman NN 8-18-4 *Np* = 246	3600 (*exp*)	–	800 (*exp*)	• G • T_A • V_{in} • I_{in} • V_{load} • I_{load} • V_{MPP} (array) • I_{MPP} (array)	• Healthy • 3 faults	• NC	NC, 4×2 mini PV array (40.3Wp)	**99.5%** (average)	• PS • OC string • Total shading	

('NC' in all columns represents non-communicated information; 'N_1-…N_i-…N_L' in 'ANN type & structure' column denotes the structure of model, i.e., N_i neurons in *i*th layer; '(*simu*) ' and '(*exp*) ' in 'Amount of data' column denotes the data obtained from simulation or experimental test, respectively; '*modu*', '*string*' or '*array*' in 'I/O setting' column represents the electrical data measured at module, string or array level, respectively; '*a*×*b*' in 'PV techno and scale' column denotes *a* parallel strings with *b* modules connected in series).

configuration and fault severities (like the cases in Ref. [49,50]). Nevertheless, in some applications, the applied model is compared on the same benchmark with either other types of ANN, or other MLTs. In these cases, the proposed ANNs have been demonstrated outperforming other ANNs or MLTs.

4) Limitations and prospects

In some cases, redundancy is found in the input feature setting, like the simultaneous use of V_{MPP}, I_{MPP} and P_{MPP}. Besides, some cases focus on limited fault types, even only one type, which is unsuitable for field application to deal with various unknown faults. However, encouragingly, some SNNs, especially MLP, have proven their efficiency (>90%) to diagnose various faults for different PV platforms with different scales. These models are promising, as they look suitable for almost all the common PV electrical faults, and can be applied in real-time health monitoring.

The use of field-measured data should be encouraged, particularly for high-scale PV platforms, which are now very common. However, where the amount of faulty data is limited, simulation data could also be adopted. However, they should be as close as possible to the data measured in the field, taking into account the uncertainties and actual measurement conditions. Also, the scope of the faults to be addressed should be broadened towards soiling, line-line fault, and faults in power electronics device.

2.1.2. Indirect FDD

Regarding the indirect applications of SNNs (i.e., output features need further interpretation for fault identification), few researches have been reported. The typical ones are presented in Table 2.

Based on Table 2, the following analysis can be conducted:

1) Integration of SNN in PV FDD

As in the previous section, MLP is still the favorite model. Similar to the prediction of PV performance [23], environmental data is used as input to predict PV electrical features that are analyzed for fault identification.

2) Detectable faults

Since all the model outputs are the electrical features, it is accordingly the faults that introduce a significant impact on the electrical features that are mainly addressed, like OC and PS.

3) Performance

The number of studies is limited compared to the direct ones, but all the reported accuracy is similarly higher than 90%.

4) Limitations and prospects

These indirect applications need an additional diagnosis module to analyze the output features. Furthermore, among these schemes, a comparison of the predicted features with the measured ones is the most common type. However, this kind of analysis could be fully realized inside one ANN by adding the measured features as inputs as it has been already done like in Refs. [37,46,52]. Therefore, the capabilities of the ANN should be fully exploited instead of complexifying the methodology.

2.2. Application of DNN for PV FDD

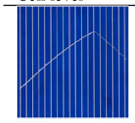
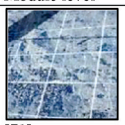
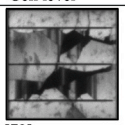
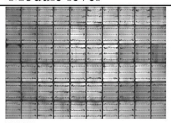
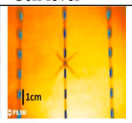
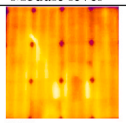
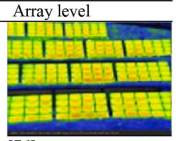
The Deep NN (DNN) applied for PV FDD has two main types, Convolutional Neural Network (CNN) and Deep Belief Network (DBN) [61]. Regarding the CNNs, various well-designed structures have been

Table 2
Details on research using indirect FDD.

Year & Ref.	ANN type, Structure & Number of parameters	Amount of data			I/O setting		Follow-up analysis	PV techno & scale	FDD accuracy	Target Fault	Notes
		Train	Validate	Test	Input	Output					
2011 [57]	MLP 4-10-6 $Np = 120$	30 (*simu*)	–	NC (*simu*)	• G • T_M • V_{MPP} (array) • I_{MPP} (array)	• 6 voltages	Use a proposed topology to analyze the 6 voltages	sc-Si, SM-55, 3×2 PV array (330Wp)	**99.9%**	• SC module (with localization)	• Number of models equals the SC cases
2015 [58]	MLP 2-15-15-1 $Np = 303$	134 (*simu*)	17 (*simu*)	17 (*simu*)	• G • T_M	• $P_{predicted}$	Compare the predicted *P* with measured one	Silero Triex U300, PV array (3.6 kWp)	**100%**	• PS • OC string • MPPT failure	
2016 [59]	MLP 2-7-12-2 $Np = 145$	2100 (*exp*)	–	900 (*exp*)	• G • T_M	• $I_{predicted}$ • $V_{predicted}$	Compare the predicted *I*, *V* with measured ones	mc-Si, BPSX150, PV module (150Wp)	NC	• PS (failure degree could be quantified)	• Input data normalized to [-1,1]
2018 [60]	MLP 2-10-1 $Np = 43$	70% (*exp*)	–	30% (*exp*)	• G • T_M	• $P_{predicted}$	Compare the predicted *P* with measured one	NC, PV array (6.2 MWp)	**90%**	• Inverter fault	

('NC' in all columns represents non-communicated information; 'N_1-…N_i-…N_L' in 'ANN type & structure' column denotes the structure of model, i.e., N_i neurons in *i*th layer; '(*simu*)' and '(*exp*)' in 'Amount of data' column denotes the data obtained from simulation or field-measurement, respectively. '*array*' in 'I/O setting' column represents the electrical data measured in array level; '*a*×*b*' in 'PV techno & scale' column denotes *a* parallel string with *b* modules connected in series.).

Table 3
Typical PV images used in DNNs for PV FDD.

Vis. images		EL images		IR images		
Cell level	Module level	Cell level	Module level	Cell level	Module level	Array level
						
[70]	[71]	[72]	[73]	[74]	[75]	[76]

practiced, such as LeNet [62], GoogLeNet [63], VGG [64], R–CNN [65], ResNet [66], AlexNet [67], Attention U-Net [68], YOLO [69]. Compared to SNN models, DNNs mainly differ in the model depth, and input data type. For the input, instead of 1-dimensional (1D) features (e.g., electrical or environmental features), DNNs deal with 2-dimensional (2D) ones, which include PV images, generated graphs or matrix.

In this section, the application of DNNs using PV image data is firstly presented. Then, the DNNs handling other types of 2D data will be investigated. At the end of each part, the applications will be discussed in the light of the four questions raised in the introduction.

2.2.1. *DNNs using PV image data*

The 2D PV data for DNNs consists of visual (Vis.), electroluminescence (EL) or thermal infrared (IR) images, as listed in Table 3. For each application, the type, structure, input type and size, pre-processing method, data partitioning and the targeted PV fault will be taken into account. All this information is summarized in Table 4 with the unfolded hierarchical structure of three examples of DNNs illustrated in Fig. 2.

Based on Table 4, the applications of DNNs using 2D PV image data are analyzed from the following aspects:

1) Integration of DNN in PV FDD

- *Model type and structure*: CNN models are the most common ones, which may be due to its popularity in image processing and pattern recognition in recent years. Among the applied CNN models, 12 out of 19 are based on the entire or partial structure of mature models (like LeNet and VGG in Fig. 2 (a) (c)), while the remaining six are thoroughly redesigned (like Fig. 2 (b)). As for the structure, the majority of DNNs are more complicated than SNN ones from both model depth and the number of parameters. However, it should be noted that more layers do not necessarily lead to higher *Np*, which can be observed from the models in Fig. 2 (b) and (c). *Np* mainly depends on the sizes of fully connected layers and the kernel, which are determined by the structure and vary from case to case.
- *Input setting*: For the origins of the images, indirect images (collected by EL or IR devices) are preferred to visual ones. EL images are the most adopted (11 use cases) as they may embed more fault information. Regarding the scope, cell level images are the most adopted (10 cases), while those at module level (6 cases) and at array level (3 cases) are relatively less investigated.
- *Data augmentation*: It is often practiced when the original dataset is not large enough or to increase the model generalization capability. In the reviewed cases, more than half (11 cases) implements this technique, with usually rotation and flip.

2) Detectable faults

At the cell level, cracking in various forms is the most frequently examined defect. At the module and network level, hot spots, delamination, soiling and interconnect failures are generally covered. A common feature of these faults is their marked presence in the adopted images. This allows the various DNNs, which are competent in pattern recognition, to identify them.

3) Performance

Overall, 13 out of 19 cases have achieved classification accuracy higher than 90%. Similarly, due to the diversity in various aspects for most applications, no further quantified global conclusion can be drawn. However, in some cases, based on the same benchmark, the proposed model has exhibited higher performance compared to other DNNs or MLTs.

4) Limitations and prospects

It is noted that more than half of the reported applications can only do binary classification, i.e., identify healthy or faulty condition. In some works, where both binary and multi-class classifications (more than 1 fault type) are conducted, the multi-class FDD accuracy is found lower than that of binary classification. This, in a sense, reflects the difficulty in precise PV faults diagnosis using image data. Besides, some models are highly complex, which could increase the computational expense. Thus, it is suggested to develop CNN model with appropriate architecture like in Ref. [73,92] to reduce unnecessary complexity. It is noteworthy that authors in Ref. [72,92] have used identical public PV image dataset [83]. Then, the comparison of different models and performance is made possible. Regarding the real-time application in large-scale power plants, DNNs also show some potential, especially via UAV-captured aerial images. However, the automatic detection of PV panels and subsequent segmentation is still a tedious task. Although some research has been dedicated to this problem [94], the robustness and precision still require further reinforcement through more field tests.

2.2.2. *DNNs using other 2D data*

In addition to PV images data, DNN models for PV FDD also process graphs or data matrix generated from 1-D features as shown in Fig. 3. For example, Lu et al. [95] and Manohar et al. [96] extract graphs from sequential *V* and *I* data. Chen et al. [97] combine I–V curve data with *G* and T_M to form up a 40×4 feature matrix. Similar I–V curve-based approaches are also applied in Ref. [98,99]. Besides, Aziz et al. adopted Continuous Wavelet Transform (CWT) [100] to generate scalograms (2-D graphs) from environmental and array electrical parameters. A similar approach is also employed in Ref. [101]. The details of these applications are presented in Table 5.

Similarly, these applications of DNNs using other 2D data are examined from the following perspectives:

1) Integration of DNN in PV FDD

In these cases, 1D features are transformed directly (like using time-series graphs) or indirectly (via specific transformation methods) to 2D features. Since these features are not real PV images, data augmentation is not generally performed.

2) Detectable faults

The common target faults are PS, OC, LLF and AF. As the inputs are

Table 4
Application of DNN models using PV image data.

Year & ref.	Model, No. of layers & Number of parameters	Input setting		Data augmentation	Amount of data				Output setting	FDD accuracy	Target fault	Notes
		Type& Size	Scope		Original	Train	Validate	Test				
2018 [77]	Adapted LeNet-5, 6 layers, *Np* = 1.1 M	EL images (100×100 pixels)	PV cell	Brightness adjustment Adding blurring Rotation	–	5120 (*aug*)	–	1000 (*aug*)	Healthy 4 faults	**98.4%** (average)	Cell crack in form: Linear Cross Flaky Broken	Trained by Nesterov Accelerated Gradient (NAG) algorithm [78]
2018 [79]	LeNet, 7 layers, *Np* = 0.06 M	EL images (NC)	PV cell	–	658	600 (*orig*)	48 (*orig*)	10 (*orig*)	Healthy 1 fault	**99%** (average)	Cell crack	Trained by Stochastic Gradient Descent (SGD) algorithm [80] Compared and outperforms GoogleNet V1 (accu: 98%)
2018 [81]	VGG-16, 21 layers, *Np* = 134.2 M	IR images (224×224 pixels)	PV array	Rotation Flip	3336	5190 (*aug*)	–	1298 (*aug*)	Healthy Defective	**75%** (best case)	Damaged module (fault type not specified)	Images captured by UAV Images manually-labeled Impact of data balance discussed
2018 [82]	Adapted VGG-16, 21 layers, *Np* = 0.38 M	EL images (120×120 pixels)	PV cell	Flip Rotation Translation Cropping	98,280	90% (*aug*)	–	10% (*aug*)	Healthy Defective	**92.3%** (balanced accu rate)	Damaged cell (different levels crack)	Impact of oversampling and augmentation discussed
2019 [72]	Adapted VGG-19, 24 layers, *Np* = 34.9 M	EL images (300×300 pixels)	PV cell	Rotation Translation Flip	2624	196,800 (*aug*)	–	656 (*orig*)	Healthy Defective	**86.6%** (for sc-Si images) **86.1%** (for mc-Si images)	Damaged cell (microcrack, insulation fault, connection fault, solder failure)	Images labeled by experts Blurry images are not excluded from dataset Public dataset [83]
2019 [84]	Mobilenet 28 layers, *Np* = 4.2 M	IR images (224×224 pixels)	PV module	Flip Rotation	798	NC (*aug*), (*orig*= 638)	–	NC (*aug*) (*orig* = 160)	Healthy 4 faults	**89.5%(average)**	Module hotspot (4 shapes of module hotspot)	Compared and outperforms VGG-16 (accu: 85.8%)
2018 [85]	Deep SolarEye (based on ResNet), *Np* = 2.0 M	Visible images (192×192 pixels) +*G*, time	PV module	Rotation Flip	NC	27,537 (*aug*)	–	18,217 (*aug*)	Soiling	**97.8%** (binary classification)	Soiling (type, impact and location)	Cleaning suggestion for soiling could also be determined
2018 [86]	DBN, 4 layers, *Np* = 4.7 M	EL images (64×64 pixels)	PV cell	–	220	200 (*orig*)	–	20 (*orig*)	Healthy 1 fault	NC	Cell crack	
2019 [87]	Fast RCNN, 28 layers, *Np* NC	IR images (NC)	PV array	–	110	83 (*orig*)	–	27 (*orig*)	Healthy Hot spot	**99.7%**	Hot spot	Images captured by UAV camera Pre-trained VGG-16 adopted to extract feature map
2020 [76]	Fast RCNN, 13 layers, *Np* NC	IR images (640×534 pixels)	PV array	–	900	800 (*orig*)	–	100 (*orig*)	Healthy Hot spot	**99.02%** 99.42% (healthy) 91.67% (hot spot)	Hot spot	Images captured by UAV camera Localization error of hot spot is 0.86 m
2020 [88]	Fast RCNN + RCNN (2 channels) *Np* NC	EL images (5232×2720 pixels)	PV module	–	1461	861 (*orig*)	–	600 (*orig*)	3 faults	**98.3%(overall)** 97.5% (Crack) 99% (Broken) 98.5% (Unsoldered)	Crack Broken Unsoldered	Fast RCNN is based on VGG-19, RCNN based on ResNet-101
2020 [89]	Attention U-net (34 layers) *Np* = 8.1 M	EL images (512×512 pixels)	PV cell	Flip Shift Rotation	828	NC (*aug*), (*orig* = 620)		NC (*aug*) (*orig* = 208)	• Healthy Defective	**100% (average)**	• Cracks Finger interruption	Impact of loss function and combination of networks discussed
2020 [90]	YOLOv3, 53 layers, *Np* NC	Visible images	PV module	–	5400	4320 (*orig*)	–	1080 (*orig*)	• Healthy • Soiling	**96.3% (average)**	• Soiling (bird drop, leaves)	Compared with Mask R–CNN (92.2%)

(*continued on next page*)

Table 4 (*continued*)

Year & ref.	Model, No. of layers & Number of parameters	Input setting		Data augmentation	Amount of data				Output setting	FDD accuracy	Target fault	Notes
		Type& Size	Scope		Original	Train	Validate	Test				
2018 [91]	Proposed CNN, 27 layers, $Np = 101.2$ M	Vis. images (256×256 pixels)	PV cell	–	21,245	15,996 (*orig*)	–	4249 (*orig*)	Healthy 6 cell defects	**94.3%** (average)	Damaged cell (broken gates, paste spot, dirty cell, thick lines, scratches, and color differences)	Impact of CNN structure, image size, training methods discussed CNN of multispectral type
2019 [71]	Proposed CNN, 9 layers, $Np = 51.7$ M	Visible images (224×224 pixels)	PV module	–	8400	5880 (*orig*)	1680 (*orig*)	840 (*orig*)	Healthy 5 faults	**97.9% (average)** 98.9% (healthy) 100% (delamination) 98.2% (soiling) 96.1% (gridline corrosion) 95.7% (snail trail) 98.5% (discoloration)	Delamination Soiling Gridline corrosion Snail trail Discoloration	Input includes manually or UAV-captured images and generated faulty images Setting of layers discussed and visualized by *t*-stochastic neighbor embedding (*t*-SNE)
2019 [92]	Proposed CNN, 9 layers, $Np = 2.5$ M	EL images (100×100 pixels)	PV cell	Flip Rotation Cropping	2624	80% (*aug*)	–	20% (*aug*)	Healthy Defective	**93.02%**	Damaged cell (microcrack, insulation fault, connection fault, solder failure)	Same dataset used in [72] Layer setting discussed Impact of augmentation discussed
2019 [73]	Proposed CNN, 5 layers, $Np = 0.2$ M	EL images (50×50 pixels)	PV cell	Flip Rotation	3550	11,360 (*aug*)	–	710 (*orig*)	Healthy 2 faults	**99.7% (average)** 100% (healthy) 100% (crack) 92% (corrosion)	Cell crack Corrosion	Impact of input size discussed Compared with SVM (accu = 99.4%) and RF (97.5%) and outperforms
2020 [93]	Proposed CNN, 9 layers, $Np = 12.9$ M	EL images (224×224 pixels)	PV cell	GAN-based model operation	1800	6400 (*aug*)	1600 (*aug*)	600 (*orig*)	Healhty 3 faults	**83% (average)** 84% (healthy) 83% (break) 82% (micro-crack) 81% (Finger-interruption)	Cell crack (3 types)	Compared and outperforms VGG-16 (accu: 66%), ResNet 50 (67%) Impact of depth, kernel size discussed
2020 [75]	Proposed CNN, 9 layers, $Np = 6.8$ M	IR images (100×100 pixels)	PV module	Flip Rotation Cropping	893	80% (*aug*) Para Run-on–>	–	20% (*aug*)	Healthy Defective	**99.23%**	Damaged module (failed interconnection, glass breakage, crack, failed/resistive soldering bonds)	Images are from IR camera and internet Model pretrained by EL images

('NC' in all columns represents non-communicated information; '(a×b)' in 'Type & Size' refers to the size of images in pixels; '(*orig*)' and '(*aug*)' in 'Amount of data' column denotes original or augmented data, respectively; The layer number indicated in 'Model' column only refers to the sum of all the inner layers).

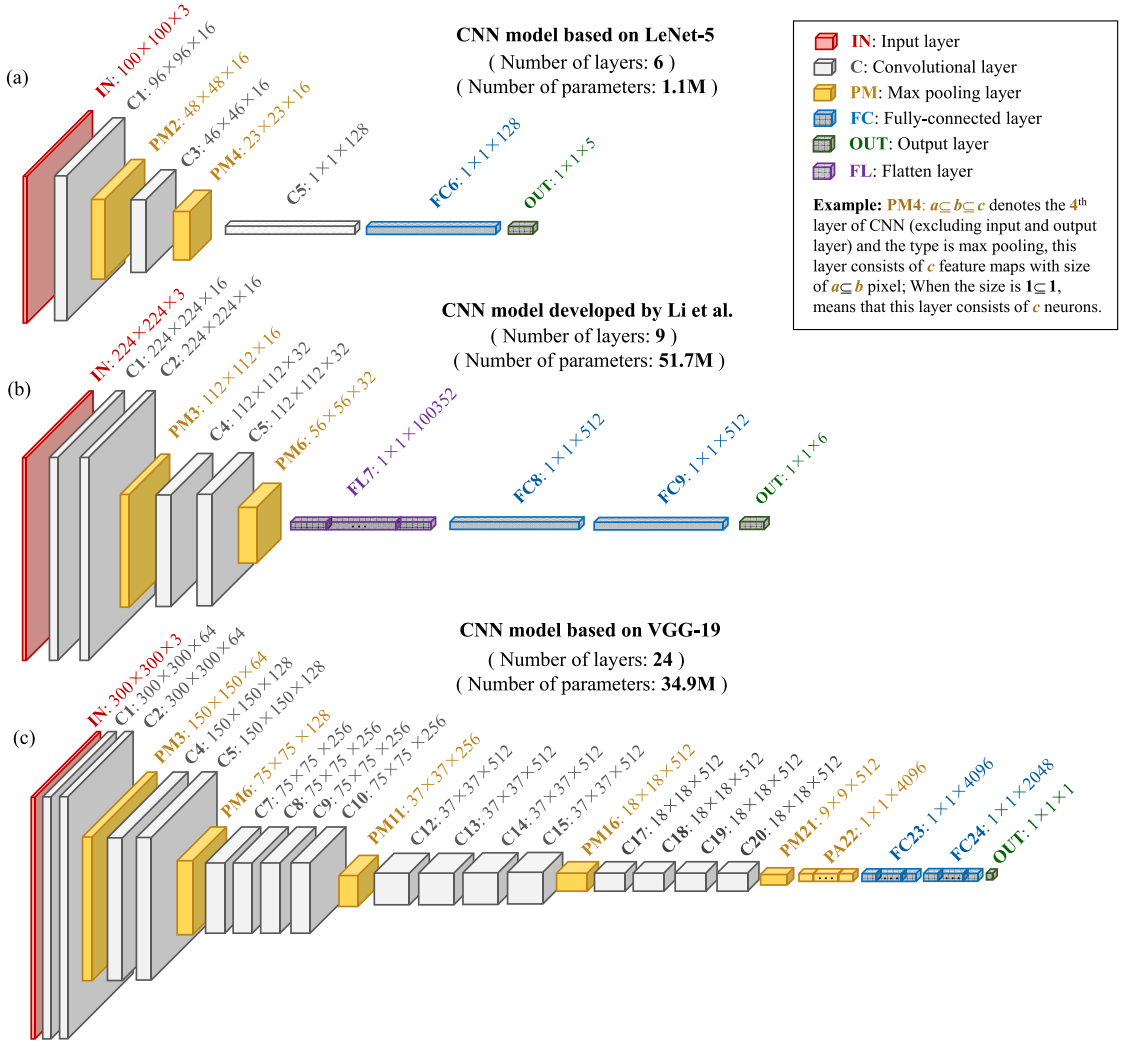


Fig. 2. Examples of CNN structures adopted for PV FDD: (a) LeNet-5 [77] (low number of layers and low number of parameters), (b) Li et al. scheme [71] (low number of layers and high number of parameters), (c) VGG-19 [72] (high number of layers and high number of parameters).

still based on the electrical or environmental 1D features, these faults are similar to the ones detected by SNNs.

3) Performance

Four out of seven cases achieved an accuracy of more than 95%, which is comparatively higher than DNNs using image data (90%). The reason behind may be fault feature separability is higher with internal variables of parameters rather than image data.

4) Limitations and prospects

Low effective information ratio is observed in some cases [95,97]. Especially in Ref. [95], the time-series graphs of current or voltage contain a large area of white zone (as shown in Fig. 3 (a)), which contributes to increase the size of the DNN input features. This problem is eased in Ref. [101,102] by adopting specific transformation techniques. The FDD schemes based on transformation from 1D to 2D have been developed to benefit from the considerable experience in image processing, with the expectation to improve the diagnosis performance. However, these schemes still require further validation and comparison with the ones using the same original 1D data and with other techniques, like the SNNs as mentioned earlier that generally have lower

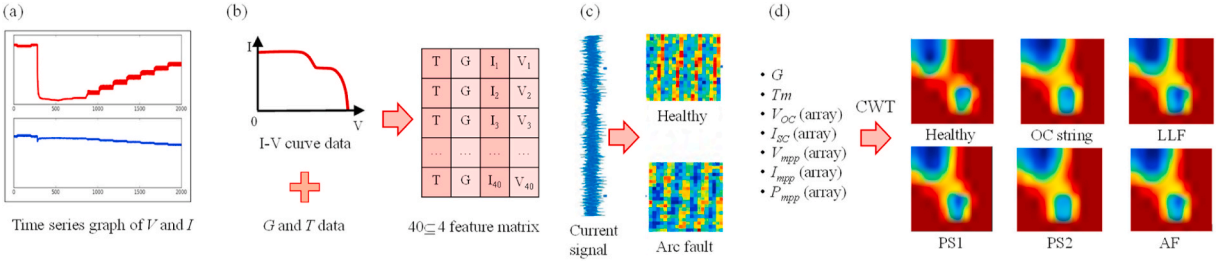


Fig. 3. Examples of other 2D input features for DNN: (a) array *I* and *V* time-series graph [95], (b) synthetic matrix [97], (c) current-generated graphs [101], (d) 1-D features-generated scalograms [102].

Table 5
Application details of DNN models using other 2D data.

Year & Ref	Model, No. of layers & Number of parameters	I/O setting		Amount of data			PV techno & scale	FDD accuracy	Target fault	Note
		Input	Output	Train	Validate	Test				
2018 [96]	SAE-DNN, *Np* NC	AC side I and V time-series graph	Healthy 6 faults	2812 (*simu*)	–	1208 (*simu*)	NC, 3×4 PV array	**100% (average)**	PS LLF (5 types)	Compared and outperforms SVM (accu: 99.2%), DT (99.3%), MLP (97.1%)
2019 [95]	Proposed CNN, 20 layers, *Np* NC	Array *I* and *V* time series graph, 299×299 pixels	Healthy 4 faults	981 (*exp*)	–	419 (*exp*)	sc-Si, GL-100 3×6 PV array (1.8 kWp)	**99.5% (average)** 99.7% (Healthy) 99.2% (SC) 98.4% (LLF) 99.9% (OC)	LLF SC module OC string (2 types)	Results obtained from 20 random tests
2019 [97]	ResNet, 18 layers, *Np* = 1863	Matrix (40×4) based on array I–V curve and T_M, G	Healthy 8 faults	15,834 (*simu*)	1858 (*simu*)	7545 (*simu*)	sc-Si, GL-100 3×6 PV array (1.8 kWp)	**99.98% (average)** 100% (Healthy, SC, PS, OC) 99.89% (Aging)	OC string PS (3 types) Aging (2 types) 1 or 2 SC module in 1 string	Unbalanced samples for each case
				1870 (*exp*)	215 (*exp*)	897 (*exp*)		**98.1% (average)** 100% (Healthy, SC) 96.68% (PS) 95.45% (OC) 95.81% (Aging)		
2020 [98]	Proposed CNN + Residual-gated recurrent unit, 7 layers, *Np* NC	Matrix based on array I–V curve and T_M, G	Healthy 10 faults	792 (*simu*)	264 (*simu*)	264 (*simu*)	sc-Si, 1×13 PV array (3.4 kWp)	**100% (average)**	PS (2 types) Aging Hybrid fault (6 types)	Impact of data missing, and anti-inference discussed
				1136 (*exp*)	379 (*exp*)	377 (*exp*)		**98.4% (average)**		
2020 [99]	Proposed CNN, 5 layers, *Np* NC	Matrix extracted from I–V curve (82×4)	Healthy 3 faults	356 (*exp*)	–	90 (*exp*)	mc-Si, 2×12 PV array	**99.8% (average)**	PS Crack Gridline fault	Compared with multi-headed NN (99.3%)
2019 [101]	Proposed GAN + CNN, 14 layers, *Np* NC	Graphs generated from array *I*, 20×20 pixels	Healthy 1 fault	24,000 (*exp*)		6000 (*exp*)	sc-Si, PV array (1.5 kWp)	**98.5% (average)** 99.3% (Healthy) 97.7% (AF)	AF	Domain adaptation applied Tested in real time FDD
2020 [102]	AlexNet (transferred), 7 layers, *Np* = 58 M	Generated scalograms, 227×227 pixels	Healthy 5 faults	2419 (*simu*)	–	1037 (*simu*)	NC, 5×3 PV array (735Wp)	Noiseless condition: **74.6% (average)** Noisy condition (noise power level of -3dB): **73.2% (average)**	AF LLF OC string PS (2 types)	Impact of training sample size and noise discussed

('(*a*×*b*)' in 'Input' column refers to the size of images in pixels; '(*simu*) ' and '(*exp*) ' in 'Amount of data' column denotes the data obtained from simulation or experimental test, respectively; '*a*×*b*' in 'PV techno and scale' column denotes *a* parallel strings with *b* modules connected in series.).

Table 6
Applications of hybridized ANN for PV FDD.

Aim of hybrid usage	Year & Ref.	ANN category, model type, structure, No. of paras	Integrated techno.	Amount of data		I/O setting		PV techno. & scale	Target faults	FDD Accuracy	Notes
				Train	Test	Input	Output				
Improve ANN algorithm	2015 [103]	*SNN*, LAPART *Np* NC	Fuzzy logic (ameliorate ANN structure)	32,400 (simu) 5760 (*exp*)	10,800 (*simu*) 1441 (*exp*)	• G • T_A • I_{MPP} (modu) • P_{MPP} (modu) • v_{WIND}	• Healthy • PS	NC, PV array (3.7 kWp)	• PS	**86% (average)** **100% (average)**	• *k*-Folds technique is adopted for the data partitioning
	2017 [104]	*SNN*, MLP (4-12-7) *Np* = 155	Particle swarm optimization (PSO) [105] (improve the global search ability)	150 (simu)	60 (*simu*)	• I_{SC} (array) • V_{OC} (array) • P_{MPP} (array) • V_{MPP} (array)	• Healthy • 6 faults	NC, 3×4 PV array (252Wp)	• PS • Aging • Temperature anomaly	**93.3% (average)**	• Compared with no PSO MLP (accu: 86.7%) • Impact of hidden layer neuron number discussed
	2018 [106]	*SNN*, MLP (6-8-7) *Np* = 125	Ada-boost classifier [107] (boost MLP performance)	1800 (*exp*)	784 (*exp*)	• G • T_A • V_{MPP} (array) • I_{MPP} (string 1–3)	• Healthy • 6 faults	sc-Si, GL-100 3×6 PV array (1.8 kWp)	• PS (2 types) • OC string (2 types) • SC module (2 types)	**97.7% (average)**	• Compared with no ada-boost MLP (accu: 96.6%)
Optimize parameter	2015 [108]	*SNN*, MLP (2-24-24-8) *Np* = 874	Genetic algorithm (GA) [109] (determine the number of neurons)	NC (*simu*)	NC (*simu*)	• V (modu) • I (modu)	• Healthy • 7 faults	NC, PV module	• PS • SC module • OC module • Aging • Charger fault • Battery fault	**97.7% (average)** 98.6% (Healthy) 96.6% (SC) 97.5% (OC) 97.6% (PS) 98.2% (Aging) 98.9% (Charger) 96.6% (Battery)	• Results under condition T_A = 33 °C, G = 380 W/m^2
	2017 [110]	*SNN*, KELM (7-NC-5) *Np* NC	Nelder-Mead Simplex (NMS) [111] (optimize kernel parameters)	3600 (*simu*) 1875 (*exp*) 4800 (*simu* +*exp*)	1200 (*simu*) 625 (*exp*) 2500 (*simu* +*exp*)	• I_{SC} (array) • V_{OC} (array) • I_{MPP} (array) • V_{MPP} (array) • n • R_S • Ep	• Healthy • 4 faults	sc-Si, GL-100, 3×6 PV array (1.8 kWp)	• PS • Aging • SC module • OC string	**100% (average)** **98.98% (average)** **98.90% (average)**	• Input data normalized by STC value • *n*: ideality factor • R_S: series resistor • *Ep*: RMSE of parameter identification
	2017 [112]	*SNN*, KELM (4-NC-4) *Np* NC	Simulated Annealing (SA) [113] (optimize model parameters)	600 (*simu*)	400 (*simu*)	• n • C_R • I_{ph} • fval	• Healthy • 3 faults	sc-Si, SM55, 3×6 PV array (990Wp)	• PS • Aging • SC module	**93.55% (Average)** 99.73% (Healthy) 91.55% (PS) 90.91% (Aging) 93.64% (SC)	• C_R: resistor coefficient • I_{ph}: photocurrent • *fval*: optimal fitness value
Pre- process or extract features	2018 [114]	*SNN*, MLP (4-12-4) *Np* = 116	Discrete wavelet transform (DWT) [115] (extract features from V_{MPP}, I_{MPP}) + Principal component analysis (PCA) (pre-process)	459 (simu)	197 (*simu*)	• 4 extracted principle components	• Healthy • 3 faults	NC, PV array (3.5 kWp)	• Module, inverter, converter fault (type not detailed)	**98.2% (average)**	• Compared with RBF NN (accu: 93.6%)
	2018 [116]	*SNN*, MLP (11-18-14) *Np* = 493	DWT based Multi-resolution analysis (MRA) [117] (extract features from converter noise signal)	7000 (*simu*)	1750 (*simu*)	• 11 signal features	14 fault locations	NC, PV array (96×17) (500 kWp)	• Cable GF • Cable LLF	**99.2% (Average)** 99.9% (Cable GF) 97.3% (Cable LLF)	• *tanh* as activation function for hidden layer and *sigmoid* for output layer • Trained by LM BP algorithm

(*continued on next page*)

Table 6 (*continued*)

Aim of hybrid usage	Year & Ref.	ANN category, model type, structure, No. of paras	Integrated techno.	Amount of data		I/O setting		PV techno. & scale	Target faults	FDD Accuracy	Notes
				Train	Test	Input	Output				
		SNN, MLP (9-18-8) $Np = 341$		4500 (*simu*)	1125 (*simu*)	• 9 signal features	8 fault locations	NC, 48×17 PV array (250Wp)	• AF • LL • Module GF	**97.4% (Average)** 96.5% (AF) 97.4% (LLF) 98.4% (Module GF)	
	2018 [118]	*SNN*, MLP 30-23-6 $Np = 887$	Multistate data processing (extract features from inverter current)	240 (*exp*)	105 (*exp*)	• Extracted features from inverter current	• Healthy • 6 faults	NC, 6×26 PV array (43.7 kWp)	• Inverter switch OC (6 locations)	**100%**	
	2018 [119]	*SNN*, PNN (4-NC-12−1)	Gaussian kernel function-based fuzzy C-means (GK-FCM) [120] (label faulty data)	492 (*exp*)	984 (*exp*)	• V_{OC} (array) • I_{SC} (array) • V_{MPP} (array) • I_{MPP} (array)	• Healthy • 11 faults	mc-Si, JKM245p, 3×13PV array (9.6 kWp)	• 2,4 or 6 SC modules in 1 string • 1 or 2 OC string • Aging (R_s = 2, 4 or 6 Ω) • 2,4 or 6 PS module in 1 string	**92.5% (Average)** 84.6% (SC) 100% (OC) 98.0% (Aging) 87.4% (PS)	• Input data normalized by STC value and G, T_M
	2019 [121]	*SNN*, MLP (5-12-8) $Np = 181$	Grey level co-occurrence matrix (GLCM) [122] (extract texture features from IR images)	NC (*orig*)	NC (*orig*)	• 5 texture features from module image subregion (20×35 pixels)	• Healthy • 7 cases of temperature anomaly	NC, PV array (169.9 kWp)	• Temperature anomaly (Hot spot and other types)	**92.8% (average)**	• *sigmoid* as activation function for output layer • Subregion divided from 1568 PV module IR images
	2019 [123]	*SNN*, PNN (48-48-1-1) $Np = 2451$	GLCM (extract features from microscope-captured cell images 1280×1024 pixels)	132 (*orig*)	240 (*orig*)	• 48 features	• Healthy • Defective	Organic PV cell	• Damaged cell (particles, bubbles, and cracks)	**95.4% (average)**	• Singular value decomposition (SVD) [124] is adopted to ease the dimensional greedy effect
	2019 [125]	*SNN*, MLP (40-100-50-30-6) $Np = 7270$	GLCM (extract texture features from module Vis. images)	5551 images in total (*orig*)		• 40 texture and statistical features	• Healthy • 5 faults	mc-Si, PV module (230Wp)	• Soiling (5 levels)	**96.3% (average)**	• Compared and outperforms kNN (accu: 93.2%), RF (94.6%), but underperforms SVM (97.5%)
	2019 [74]	*DNN*, GoogLeNet (21 layers) $Np = 6$ M	PCA (extract data from image sequences)	720 (*orig*)	720 (*orig*)	•Extracted cell IR image (224×224 pixels)	• 6 faults	c-Si, PV cell	• Crack • Hot spot • Scratch • Broken edge • Surface impurity • Large area damage	**99.7% (average)**	• Compared and outperforms VGG-16 (accu:99.1%), LeNet-5 (98.9%), SVM (98%) and MLP (97.3%) • Pre-process techniques compared
	2019 [126]	*DNN*, Proposed CNN, (9 layers) $Np = 0.37$ M	Sliding window scan (extract regions from cell EL image 943×923 pixels)	NC (*aug*), (*orig*= 499)	NC (*aug*) (*orig* = 43)	• Regions (128×128 pixels)	• Defect probability of regions	NC, PV cell	• Defective cell	**88.4% (average)**	• 4 types of CNN structure tested
	2019 [127]	*DNN*, ResNet (50 layers)	Denoising CNN (DnCNN) [128] (extract soiling features)	3520 (*aug*)	880 (*aug*)	• Extracted dust layers from visible module images (244×244 pixels)	• Soiling rate	mc-Si, CSUN260–60P, PV module (260Wp)	• Soiling	**F1: 90%** **RMSE: 0.69**	• Images augmented by rotation and flip • Original image amount: 550 • Compared and outperforms other depth of ResNet

(*continued on next page*)

Table 6 (*continued*)

Aim of hybrid usage	Year & Ref.	ANN category, model type, structure, No. of paras	Integrated techno.	Amount of data		I/O setting		PV techno. & scale	Target faults	FDD Accuracy	Notes
				Train	Test	Input	Output				
Post- process	2018 [129]	*DNN*, VGG-16 (21 layers)	SVM (classify the features extracted from pretrained model)	2400 (*aug*)	600 (*aug*)	•Visible cell images (224×224 pixels)	• Feature vector of 4096 dimension	NC, PV array	• Delamination • Discoloration • Glass breakage • Soiling • Snail tracks	**90.2% (Binary classification)** 78.9% (Delamination) 77.3% (Discoloration) 76.9% (Breakage) 84.4% (Soiling) 77.8% (Snail tracks)	• Best results obtained on the model pre-trained by ImageNet • Images captured by UAV drone
	2019 [130]	*DNN*, *Proposed CNN, 9 layers,* *Np* = 25.7 M	SVM (classify the features extracted from pretrained model)	NC (*aug*)	NC (*aug*)	•Visible module images (224×224 pixels)	• Feature vector of 512 dimension	NC, PV module	• Glass breakage • Delamination • Soiling • Corrosion • Snail trail • Discoloration	**98.1% (average)** 100% (healthy) 96% (glass) 98% (delamination) 96% (soiling) 97% (corrosion) 98% (snail trail) 100% (discolor)	• Transfer learning based on [71] • Setting of transfer discussed • Compared with AlexNet (accu: 85.9%), VGG16 (68%) and outperforms
	2020 [131]	*DNN*, VGG-16 (21 layers) *Np* = 134.2 M	Low rank matrix recovery (LRMR) [132] (decompose output features for classification)	1372 (*orig*)	1372 (*orig*)	• Vis.+EL cell images (224×224 pixels)	• Healthy • Cell crack	mc-Si, sc-Si, PV cell	• Cracks	**F1 :46.8%**	• VGG pre-trained on ImageNet Same EL dataset used in [72] • Compared and outperform non LRMR (F1: 42.6%)

('NC' in all columns represents non-communicated information; 'N_1-…N_i-…N_L' in 'ANN type & structure' column denotes the structure of model, i.e., N_i neurons in ith layer; '(*simu*) ' and '(*exp*) ' in 'Amount of data' column denotes the data obtained from simulation or experimental test, respectively, '*orig*' and '*aug*' denotes the used images or the images where the input features are extracted are original or augmented, respectively; '*modu*' and '*array*' in 'I/O setting' column represents the electrical data measured in module level or array level, respectively; 'a×b' in 'PV techno & scale' column denotes a parallel strings with b modules connected in series.).

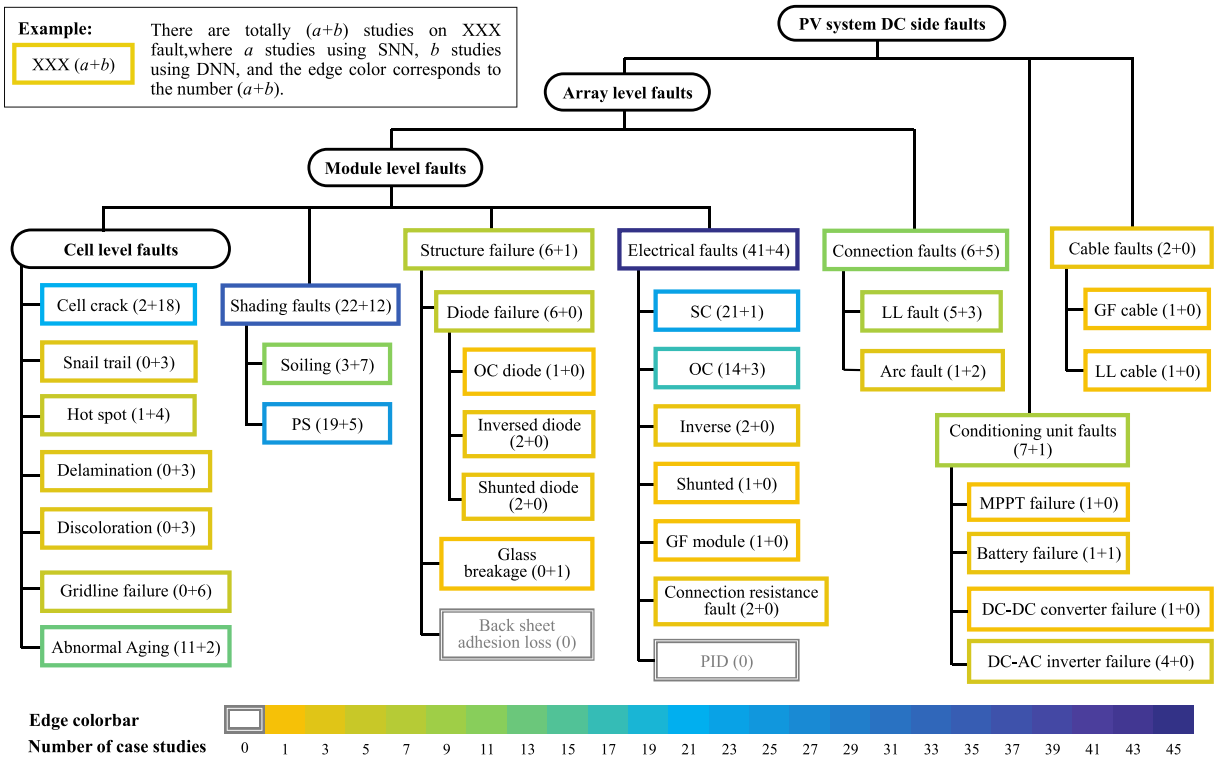


Fig. 4. Statistical chart of PV faults detection using ANN techniques.

complexity.

2.3. Hybridization of ANN for PV FDD

In order to enhance the ANN performance, several researchers have combined ANN with other techniques for different objectives, like:

- *Extractionor pre-processing of input features*,
- Improvement of ANN algorithm,
- Optimization of ANN parameters,
- Post-processing of output features.

According to these specific objectives, the applications of the hybridized techniques are summarized in Table 6.

As it has been done for the other techniques, hybridized ANNs are analyzed in light of the questions raised in the introduction:

1) Integration of hybridized ANN in PV FDD
   - *Integrated technology & usage*: Among the studied hybridization applications, extraction and pre-processing of input features are the most common objectives, covering both SNN and DNN models. Using DWT for electrical signals, and GLCM to extract features from images are the typical schemes. These techniques serve to eliminate useless information and improve the quality of input, which is essential to the performance of one ANN model. Besides, some post-processing techniques are also applied, generally in DNNs, to classify high-dimension output features. In other cases, the techniques assist in mitigating the inherent limits of the ANN technique, like the tricky procedure of parameter optimization, the time-consuming labelling work, and the imperfect learning algorithm.
   - *ANN model*: Both SNNs (13 cases) and DNNs (6 cases) have been practiced for hybridization. Among SNNs, MLP is still the most popular. Among DNNs, CNNs with various structures are more considered. The favored model type and other aspects of the model configuration are similar to the aforementioned observations in the single applications of SNNs or DNNs.
2) Detectable faults:

When based on 1D features (i.e., before pre-processing), the typical target faults include PS, OC, SC and LLF. For those dealing with 2D features, defects at cell level and soiling fault are more investigated.

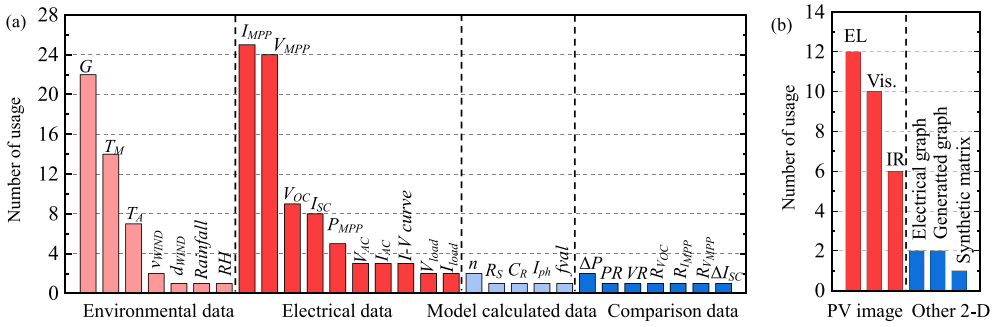


Fig. 5. ANN input data for PV FDD, (a) 1D data, (b) 2D data.

Table 7
Comparison of key features of 3 scenarios of ANN for PV FDD.

Application scenarios	Common types	Common input data	Distribution of input data	Common pre-processing operations	Common target PV faults	Advantages	Disadvantages
Application of SNN models	• MLP • RBF • PNN • ENN • ELM	• Electrical data • Environmental data	• 57.6% cases use simulation data, 42.4% use experimental data • 92.6% electrical data are at string or array level, 7.4% are at module level	• Normalization • Scaling by STC value or environmental data	• Electrical faults (like SC, OC, LLF) • PS • Aging	• Simple structure • Easy access of a large number of data (simulation type)	• Weak in processing 2D data
Application of DNN models	• CNN • ResNet • DBN	• PV image data • Generated graph • Synthetic matrix	• 55.6% case use augmented data, 44.4% original data • 55.6% input images at cell level, 33.3% at module level, 11.1% at array level	• Image resizing and segmentation • Data augmentation	• Permanent visible faults (like cell crack, snail trails)	• Efficient in image processing • Dataset for pre-training available on the Internet • Mature models available on the Internet for reference • Shared weights (for CNN and ResNet)	• Require a large number of labeled 2-D data • High computational complexity • Prone to gradient vanishing
Hybrid application	• For parameter optimization • For pre-processing • For post processing	Depend on SNN or DNN				• Mitigate the inherent limits of ANN technique • Provide input features with richer information • Ameliorate the FDD performance	• Effort needed to realize efficient integration

3) Performance:

Overall, the majority of applications (17 out of 19) achieve accuracy better than 90%. In some cases, the comparison showed that hybridization improves classification performance.

4) Limitations and prospects:

In some cases, the hybridization could increase the complexity of the FDD scheme. However, it is worthy to accept this cost if a significant amelioration is achieved. Therefore, the performance comparison is necessary to validate the combined techniques. However, this is not always reported. Nevertheless, globally, hybridization is promising, not only to improve the accuracy, but also to make possible the transformation of features from 2D to1D. Like in Ref. [115], 1D features extracted from 2D images are used as input. On one side, this transformation reduces the input features dimension, and, on the other side, allows the use of simple SNN for the diagnosis instead of DNN. This approach can significantly simplify the tuning and training process of the NN.

3. Discussion

For a better understanding of the previously described applications (73 cases), a discussion from a statistical perspective is presented in this section to provide an answer to each of the four questions raised in the introduction.

3.1. Which are the PV faults detectable by ANN?

All the common faults [8] are firstly presented in the form of 'PV DC-side fault tree' as in Fig. 4. It is based on 4 structural levels, i.e., PV cell, module, array and DC side level. The faults detected by the ANN technique are then summarized and marked as 'leaves' of the 'tree', where the number of studies using SNN or DNN is presented after the fault name, and the total number is expressed by the nuance of the block edge color (darker represents more studies).

Overall, the presented ANN techniques with different configurations can detect almost all the common faults on the PV system DC side. However, some faults like PID (almost eliminated thanks to technological improvement) and back-sheet adhesion loss are uncovered. This may be due to the detection's difficulty (limited input data, weak fault signature) or current little research interest in this context. From a quantitative aspect, short circuit module, partial shading, open circuit module and abnormal aging are the most studied ones because of their stronger signature and easy reproducibility in simulation or field tests.

Regarding the applied ANN methodology, almost all the permanently visible faults (cell crack, delamination, discoloration) are usually detected with DNN models. The other ones, like shading, structure failure and electrical faults, are detected with SNN models. This difference may mainly come from the characteristics of the fault. Indeed, permanently visible faults generally introduce a more significant impact on PV images data rather than on 1D data. Thus, DNN models that can conduct efficient image processing (using 2D data) are preferred while SNN models seem more suitable for the other faults (using 1D data).

3.2. What are the performance?

As we have seen previously, it is difficult to give a clear answer with regard to the various applications developed on different platforms. Nevertheless, on the basis of the results summarized in Tables 1 and 2 and 4–6, some partial conclusions can be drawn:

- Most of the proposed ANNs can achieve classification accuracy higher than 90%, while in other cases, the relatively less satisfying performance is supposed mainly due to the low separability of the

adopted input features or the improper type or structure of the developed models.

- Hybrid models generally perform better than the original ones with the average reported improvement of accuracy as 3.9% relying on the integrated techniques.
- In some applications, using the same benchmark, the proposed ANN outperforms other MLTs (like DT, RF, SVM, etc.) with the average reported improvement of accuracy as 3.8%. However, it should be noted that this shows only a general trend and does not necessarily mean that the ANN always outperforms other MLTs. These results are valid in the cases presented and depend on how these candidate techniques are developed and parameterized.

3.3. How is ANN integrated in PV FDD?

3.3.1. Source of dataset

The input 1D data is obtained at 57.6% from simulation and 42.4% from field measures among the reported training, validation and test dataset. The popularity of simulation data may be owing to its good controllability of test conditions (e.g., independent control of G and T) and low acquisition cost. However, the approximation to real PV conditions is restricted. On the opposite, despite field data reflects the operational condition, its wide application is still limited because of weak condition controllability (weather conditions, season, etc.) and measurement issues (accessibility, sensors). Therefore, the selection of data sources is a trade-off among all the aforementioned factors. Whereas, it is still recommended to use at least the field data as the test dataset to evaluate model performance under real conditions.

2D data includes PV image, generated graph or matrix. With regard to image dataset, except very few cases that employ generated PV faulty images, the majority of the adopted dataset is captured in the field. In recent years, taking benefit from the rapid development of drone technology, UAV with an embedded camera has been widely used for remote inspection of PV power plants. Abundant EL or IR images have been gathered. Nevertheless, the quantity of the original dataset is always inadequate to fulfill a fine-tuning and may consequently introduce the underfitting problem (i.e., a model cannot capture any trend [133]). A common solution is to perform rotation, flip, adding blurry, adjusting illumination or other operations to enrich the original dataset. These transformations are found convenient and efficient to enhance the generalization ability of the DNNs. As for the generated graph or matrix that are based on 1D features, the first results are promising but still require further validation of the necessity of the 1D-to-2D FDD scheme.

3.3.2. Data pre-processing

Since 1D values lie usually in different ranges, e.g. irradiations range from 0 to 1300 W/m^2 while the temperature varies from −10 to 80 °C (in the reviewed applications), they are usually normalized before being introduced into the ANN. The results show a significant decrease in the iteration process, especially when the features are centered and standardized.

For 2D data, the unification of image size and digital pre-processing (like graying operation, RGB separation) are also necessary.

3.3.3. Type of data

1D data is presented in Fig. 5 (a) with 4 categories (environmental, electrical, model-calculated, and compared data (ratio between 2 parameters)) and 2D data is presented in Fig. 5 (b).

For 1D data, environmental and electrical measures are the most frequent. Among single data type, V_{MPP} and I_{MPP} are the most selected ones, as they reflect the MPP changes due to various faults. G (specifically, plane-of-array irradiance) and T_M (backplane module temperature), which are almost always measured have gained similar acceptance because of their impact on PV performance. V_{OC} and I_{SC} are also employed but to a lesser extent probably due to the measurement difficulty when the PV system is operating. The model-calculated and compared data are up to now less used probably due to PV model accuracy and calculation complexity for parameter extraction.

For 2D data, EL and visual images are preferred to IR and other data types as input features to DNN models. For IR technology, the cost and complexity of the equipment is still an issue, and moreover the operating point and the environmental conditions (outside temperature, wind velocity, …) make it difficult to interpret and process the images. The other 2D data types are just beginning to be used, and their application will grow with new 1D-to-2D transformation techniques.

3.3.5. Parameter configuration

For each ANN model, there are several key configurations, the number of hidden layers, the activation function, the loss function, the learning rate, the combination of convolutional and pooling layers (for CNN), all of which are troublesome to be directly determined. Therefore, comparative studies with different settings are crucial to finding an optimal configuration, like the reported cases in Ref. [38,47,71].

Besides, since CNN models are commonly more complicated than SNN ones, there are some open-access models pre-trained with standard image datasets (like ImageNet [137]). These models could be adopted and get further trained by PV image dataset, which may save the training time and promote the PV FDD performance, as the application in Refs. [129]. However, these models, generally with complex and deep structures, may introduce unnecessary redundancy for the application to PV FDD. Consequently, under the circumstances with limited computing capability or using a large number of samples, the structure of these models should also be thoroughly re-examined.

3.3.6. Summary of applications

The three reviewed application scenarios using SNN, DNN or hybrid methods, differ in various aspects, all of which have been investigated and discussed before. Now, the key features of these scenarios are

Table 8
Public PV fault datasets.

Ref.	Data type	Amount of data	Fault type	Lien
[83]	EL images (PV cell, 300×300 pixels)	2426	• Defective cell	https://github.com/zaebayern/elpv-dataset
[73]	EL images (PV cell, 250×250 pixels)	1031	• Cracks • Corrosion	https://osf.io/v6pwe/
[85]	Vis. Images (PV module, 192×192 pixels)	45,754	• Soiling	https://deep-solar-eye.github.io/
[52]	1D data (environmental and electrical parameters of PV array)	3000 sets	• LLF • OC string	https://github.com/benjamin-2044/PV_fault_Python/tree/master

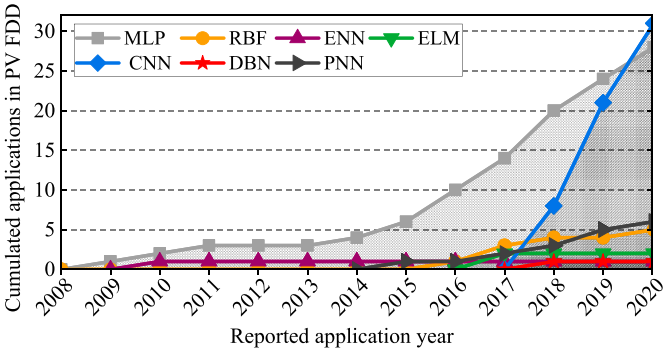


Fig. 6. Cumulative publications number of PV FDD using different ANN models (until July 2020).

summarized and compared (Table 7). The pros and cons are also presented.

3.4. What are the challenges and prospects of ANN for PV FDD?

For SNNs, DNNs, and hybrid applications, their corresponding limitations and prospects have been detailed. Here, some common points will be discussed as follows:

3.4.1. Challenges

- Model configuration

Generally, the model configuration is a tough and tedious task. Various aspects like model structure (layer number, order and size), learning algorithm, loss function, activation function, need to be determined. Regarding the application in PV FDD, when using certain types of PV features and targeting specific types of faults, there should be a commonality in the best-performing models. Under these cases, instructional strategies for the model configuration are expected to be given, however, this is not available yet and thus requires further exploration.

- Public PV fault database

In the reviewed literature, the majority of the research relies on their own-developed dataset that makes the comparison of different proposed models and performance nearly impossible. A public database, containing 1D or 2D features for common PV faults and technologies, is in desperate need. This lack is particularly real for 2D image data, since it is time-consuming and expensive to collect large amounts of PV images and do the corresponding labeling work. Up to now, in the reviewed work, some shared datasets (both 1D and 2D type) are presented in Table 8. However, the amount and covered fault types still require further enrichment and integration.

3.4.2. Prospects

- Other candidate models to be evaluated

Considering PV FDD is an interdisciplinary issue between PV technology and health monitoring, it is essential for researchers to pay additional attention to the development of related FDD techniques. For SNNs, some classical models are still not fully explored, like various autoencoder [138] (sparse, denoising and contractive type), modular neural network [139] and many variants of recurrent neural network [140] (e.g., fully recurrent, Hopfield, bidirectional type). For DNNs, driven by the great research interest in pattern recognition and deep learning, models with new structures spring up, like EfficientNet [141], DenseNet [142], SqueezeNet [143], ShuffleNet [144], etc., all of which could be evaluated for PV FDD.

- Different types of input data

Since CNN appears as an efficient tool for image classification, it can be applied to PV images captured by other techniques (e.g., pulse thermography, lock-in thermography and UV fluorescence [142]). This method may permit the identification of more PV fault types (like PID) not covered by traditional ANN models. Besides, with feature-extraction or transformation implemented, various novel 2D-to-1D or 1D-to-2D features could be adopted.

- Hybrid methods

Although ANN techniques hold some inherent limits, hybrid methods with feature-extraction approaches, parameter-optimization methods or MLTs are likely to mitigate the problems. In addition, there are various other proved-efficient data-driven methods, e.g., T-test [145], Linear Discriminant Analysis (LDA) [146], clustering [147], which can also be combined with ANN. Hence, hybridized ANN for PV FDD would become a promising research topic.

- Real-time health monitoring

Another merit of ANN is its rapid decision making. With a well-trained model, high-precision real-time health monitoring for PV arrays is made possible. Besides, based on the monitoring results, protection functions could be enabled, like the ANN-integrated relay operation of PV microgrid in Refs. [96,101].

- Experiences from other fields

ANN technique is commonly applied, not just in PV FDD, but also in the health monitoring of other systems, like for wind turbines [148], gas turbines [149], rolling bearing [150], etc. Consequently, the practice of ANN technique in these related fields, like the configuration of the model, pre- or post-processing methods, and hybridization, may provide useful experiences for the application in PV FDD.

3.3.4. Applied ANN models

All the reported model types with their cumulative number of publications are presented in Fig. 6. It is observed that, among the various ANN types, MLP and CNN are the most popular ones. This, to a certain degree, shows their efficiency and adaptability. In the meantime, other models like RBF and PNN have also been exploited for a limited number of applications. However, it should be noted that research interest in CNN models has grown rapidly since 2017, making it possible to envisage its development in the coming years.

On the other hand, it is also necessary, for each model, to keep an eye on the time gap between the theoretical development and its first application as presented in Fig. 6.[3]

Regarding the SNN models in Fig. 7 (a), ample time gaps are observed for almost all the models. This is particularly evident for RBF, PNN and MLP with more than 20 years. However, ENN is an exception with only a 7-year time gap probably because its creator has conducted the application. This phenomenon of large theory-to-application time gap may be due to three reasons: firstly, based on the literature [134–136], it has been observed that research on PV FDD started in the 1980s with conventional methods (power loss or I–V curve analysis); secondly, the increasing availability of data collected from PV power plant; finally the growing concern of operators to improve the energy efficiency of their plants.

As for the DNN models shown in Fig. 7 (b), large time-gaps are also identified for the LeNet and DBN models. Nevertheless, the time-gap displays a decreasing trend for the newly developed models, especially for VGG, GoogLeNet and ResNet whose applications have started in 2018. This evolution may be owing to the gradual maturity and dissemination of image processing and pattern recognition.

4. Conclusion

In this paper, a literature overview for the application of Artificial neural networks for PV fault detection and diagnosis is performed. Artificial neural networks have been proved efficient in the detection and diagnosis task for nearly all the common PV faults, including both the electrical faults (reflected in 1-dimension features, commonly dealt with shallow neural networks) and permanent visible faults (reflected in 2-dimension features, commonly dealt with deep neural networks). The

[3] **The MLP model discussed here refers to the MLP integrated with BP learning algorithm, which was firstly reported in 1986** [24].

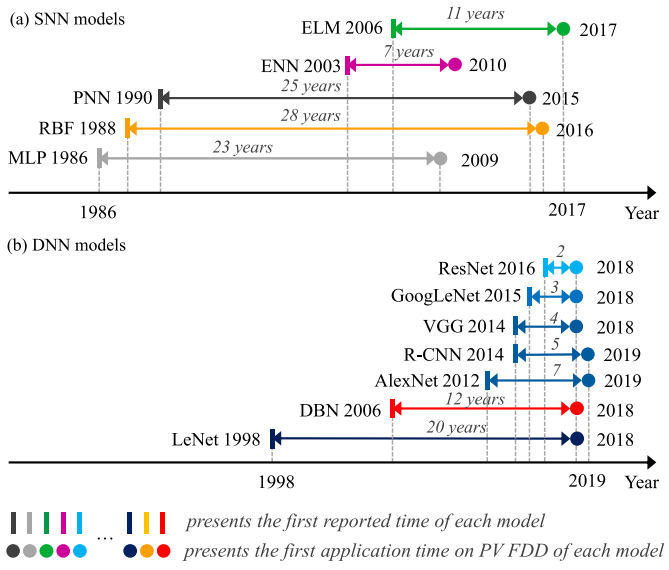


Fig. 7. Theory-to-application time gaps of models in PV FDD (a) for SNN models, (b) for DNN models.

reported classification success rate is higher than 90%. Besides, in some cases, performance comparisons with other machine learning techniques have shown the superiority of the proposed model. As for the model type, *Multi-layer perception neural network* and *Convolutional neural network* are identified as the most commonly adopted shallow and deep neural network, respectively. They can be adopted as starting models for future research.

Based on the detailed analyses of use cases, some challenges were identified. The most common ones are the difficulty in configuring the model, and the low availability of an open database on PV system failures. The latter is particularly challenging for deep neural network, as the number of faulty PV images is important for learning. It is therefore recommended that research groups or operators share as much as possible their databases with 1 and 2-dimension characteristics for healthy and different fault conditions. This database cannot only improve the generalizability of the proposed models, but will also facilitate comparisons.

Also, some prospects have been highlighted. On one side, with the rapid development of deep learning and PV technologies, more types of models and potential input features (including transformed features between 1 and 2 dimension) deserve to be explored. On the other side, the hybridization with other techniques and the application in online health monitoring are worth developing. This review is expected to be useful both for first users and experts in the hot topic of PV plant health monitoring.

Credit author statement

Li, B: Conceptualization, Methodology, Investigation, Formal analysis, writing-original, review and editing. Delpha, C: Supervision, Conceptualization, Methodology, Investigation, Formal analysis, writing-original, review and editing. Diallo, D: Supervision, Conceptualization, Methodology, Investigation, Formal analysis, writing-original, review and editing. Migan-Dubois, A: Supervision, Conceptualization, Methodology, Investigation, Formal analysis, writing-original, review and editing

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to thank the China Scholarship Council for Ph.D. funding.

References

[1] Peinado Gonzalo A, Pliego Marugán A, García Márquez FP. Survey of maintenance management for photovoltaic power systems. Renew Sustain Energy Rev 2020;134:110347.
[2] Santhakumari M, Sagar N. A review of the environmental factors degrading the performance of silicon wafer-based photovoltaic modules: failure detection methods and essential mitigation techniques. Renew Sustain Energy Rev 2019; 110:83–100.
[3] Mellit A, Tina GM, Kalogirou SA. Fault detection and diagnosis methods for photovoltaic systems: a review. Renew Sustain Energy Rev 2018;91:1–17.
[4] Tsanakas JA, Ha L, Buerhop C. Faults and infrared thermographic diagnosis in operating c-Si photovoltaic modules: a review of research and future challenges. Renew Sustain Energy Rev 2016;62:695–709.
[5] Hernández-Callejo L, Gallardo-Saavedra S, Alonso-Gómez V. A review of photovoltaic systems: design, operation and maintenance. Sol Energy 2019;188: 426–40.
[6] Correa-Betanzo C, Calleja H, Aguilar C, Lopez-Nunez AR, Rodriguez E. Photovoltaic-based DC microgrid with partial shading and fault tolerance. J Mod Power Syst Clean Energy 2019;7:340–9.
[7] Firth SK, Lomas KJ, Rees SJ. A simple model of PV system performance and its use in fault detection. Sol Energy 2010;84:624–35.
[8] Pillai DS, Rajasekar N. A comprehensive review on protection challenges and fault diagnosis in PV systems. Renew Sustain Energy Rev 2018;91:18–40.
[9] Livera A, Theristis M, Makrides G, Georghiou GE. Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems. Renew Energy 2019;133:126–43.
[10] Youssef A, El-Telbany M, Zekry A. The role of artificial intelligence in photovoltaic systems design and control: a review. Renew Sustain Energy Rev 2017;78: 72–9.
[11] Yi Z, Etemadi AH. Fault Detection for photovoltaic systems based on multi-resolution signal decomposition and fuzzy inference systems. IEEE Trans Smart Grid 2017;8:1274–83.
[12] Harrou F, Dairi A, Taghezouit B, Sun Y. An unsupervised monitoring procedure for detecting anomalies in photovoltaic systems using a one-class Support Vector Machine. Sol Energy 2019;179:48–58.
[13] Madeti SR, Singh SN. Modeling of PV system based on experimental data for fault detection using kNN method. Sol Energy 2018;173:139–51.
[14] Benkercha R, Moulahoum S. Fault detection and diagnosis based on C4.5 decision tree algorithm for grid connected PV system. Sol Energy 2018;173:610–34.
[15] Heinrich M, Meunier S, Samé A, Quéval L, Darga A, Oukhellou L, Multon B. Detection of cleaning interventions on photovoltaic modules with machine learning. Appl Energy 2020;263:114642.
[16] Delalleau O, Bengio Y. Shallow vs. deep sum-product networks. Advances in Neural Information Processing Systems, Granada, Spain 2011:666–74.
[17] Nelson D, Wang J. Introduction to artificial neural systems. Neurocomputing 1992;4:328–30.
[18] Elsheikh AH, Sharshir SW, Abd Elaziz M, Kabeel AE, Guilan W, Haiou Z. Modeling of solar energy systems using artificial neural network: a comprehensive review. Sol Energy 2019;180:622–39.
[19] Mellit A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: a review. Prog Energy Combust Sci 2008;34:574–632.
[20] Firth SK, Lomas KJ, Rees SJ. A simple model of PV system performance and its use in fault detection. Sol Energy 2010;84:624–35.
[21] Qazi A, Fayaz H, Wadi A, Raj RG, Rahim NA, Khan WA. The artificial neural network for solar radiation prediction and designing solar systems: a systematic literature review. J Clean Prod 2015;104:1–12.
[22] Mellit A, Kalogirou SA, Hontoria L, Shaari S. Artificial intelligence techniques for sizing photovoltaic systems: a review. Renew Sustain Energy Rev 2009;13: 406–19.
[23] Ghritlahre HK, Prasad RK. Application of ANN technique to predict the performance of solar collector systems - a review. Renew Sustain Energy Rev 2018;84:75–88.
[24] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. Nature 1986;323:533–6.
[25] Park J, Sandberg IW. Universal approximation using radial-basis-function networks. Neural Comput 1991;3:246–57.
[26] Specht DF. Probabilistic neural networks. Neural Network 1990;3:109–18.
[27] Wang MH, Hung CP. Extension neural network and its applications. Neural Network 2003;16:779–84.
[28] Huang G Bin, Zhu QY, Siew CK. Extreme learning machine: theory and applications. Neurocomputing 2006;70:489–501.
[29] Elman JL. Finding structure in time. Cognit Sci 1990;14:179–211.
[30] Alexandridis AK, Zapranis AD. Wavelet neural networks: a practical guide. Neural Network 2013;42:1–27.
[31] Wu Y, Lan Q, Sun Y. Application of BP neural network fault diagnosis in solar photovoltaic system. 2009 IEEE International Conference on mechatronics and automation (ICMA), Changchun, China; 9-12 aug. 2009, p. 2581–2585.

[32] Zhang WJ, Ge Q, Huang CY. The research of photovoltaic array intelligent fault diagnosis based on the BP neural network. Adv Mater Res 2014;936:2201–6.
[33] Lin H, Chen Z, Wu L, Lin P, Cheng S. On-line monitoring and fault diagnosis of PV array based on BP neural network optimized by genetic algorithm. Lect Notes Comput Sci 2015;9426:102–12.
[34] Chunlai L, Xianshuang Z, Gudake. A survey of online fault diagnosis for PV module based on BP neural network. In: 2016 International Conference on smart city and systems engineering (ICSCSE), hunan, China; 25-26 Nov; 2016. p. 483–6.
[35] Salem F, Awadallah MA. Detection and assessment of partial shading in photovoltaic arrays. J Electr Syst Inf Technol 2016;3:23–32.
[36] Chine W, Mellit A, Lughi V, Malek A, Sulligoi G, Massi Pavan A. A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks. Renew Energy 2016;90:501–12.
[37] Khelil CK, Kara K, Chouder A. Fault detection of the photovoltaic system by artificial neural networks. In: 4th International Conference on green energy and environmental engineering (GEEE-2017), sousse, Tunisia; Apr. 2017. p. 22–4.
[38] Chine W, Mellit A. ANN-based fault diagnosis technique for photovoltaic stings. In: 2017 5th International Conference on electrical engineering - Boumerdes (ICEE-B), Boumerdes, Algeria; 29-31 Oct; 2017. p. 1–4.
[39] Laamami S, Benhamed M, Sbita L. Artificial neural network-based fault detection and classification for photovoltaic system. In: 2017 International Conference on green energy Conversion systems (GECS), Hammamet, Tunisia; 23-25 Mar; 2017. p. 1–7.
[40] Sabri N, Tlemcani A, Chouder A. Faults diagnosis in stand-alone photovoltaic system using artificial neural network. In: 2018 6th International Conference on control engineering and information technology (CEIT), Istanbul, Turkey; 25-27 Oct; 2018. p. 1–6.
[41] Da Costa CH, Moritz GL, Lazzaretti AE, Mulinari BM, Ancelmo HC, Rodrigues MP, Oroski E, De Goes RE. A comparison of machine learning-based methods for fault classification in photovoltaic systems. In: 2019 IEEE PES Conference on Innovative smart grid technologies, gramado, Brazil; 15-18 sept; 2019. p. 1–6.
[42] Laarabi B, May Tzuc O, Dahlioui D, Bassam A, Flota-Bañuelos M, Barhdadi A. Artificial neural network modeling and sensitivity analysis for soiling effects on photovoltaic panels in Morocco. Superlattice Microst 2019;127:139–50.
[43] Ul-Haq A, Sindi HF, Gul S, Jalal M. Modeling and fault categorization in thin-film and crystalline PV arrays through multilayer neural network algorithm. IEEE Access 2020;8:102235–55.
[44] Andrei N. Scaled conjugate gradient algorithms for unconstrained optimization. Comput Optim Appl 2007;38:401–16.
[45] Li K, Zhao S, Wang Y. A planar location method for DC arc faults using dual radiation detection points and DANN. IEEE Trans Instrum Meas 2020;69: 5478–87.
[46] Pahwa K, Sharma M, Saggu MS, Mandpura AK. Performance evaluation of machine learning techniques for fault detection and classification in PV array systems. In: 2020 7th International Conference on signal processing and integrated networks (SPIN), Noida, India; 27-28 Feb; 2020. p. 791–6.
[47] Dhimish M, Holmes V, Mehrdadi B, Dales M. Comparing Mamdani Sugeno fuzzy logic and RBF ANN network for PV fault detection. Renew Energy 2018;117: 257–74.
[48] Hussain M, Dhimish M, Titarenko S, Mather P. Artificial neural network based photovoltaic fault detection algorithm integrating two bi-directional input parameters. Renew Energy 2020;155:1272–92.
[49] Akram MN, Lotfifard S. Modeling and health monitoring of DC side of photovoltaic array. IEEE Trans Sustain Energy 2015;6:1245–53.
[50] Garoudja E, Chouder A, Kara K, Silvestre S. An enhanced machine learning based approach for failures detection and diagnosis of PV systems. Energy Convers Manag 2017;151:496–513.
[51] Wang XX, Dong L, Liu SY, Hao Y, Wang B. A fault classification method of photovoltaic array based on probabilistic neural network. In: Proceedings of the 31st Chinese control and decision Conference (CCDC), Nanchang, China; 3-5 Jun; 2019. p. 5260–5.
[52] Basnet B, Chun H, Bang J. An intelligent Fault Detection model for fault detection in photovoltaic systems. J Sensors 2020;2020:1–11.
[53] Duan K, Keerthi SS, Poo AN. Evaluation of simple performance measures for tuning SVM hyperparameters. Neurocomputing 2003;51:41–59.
[54] Chao K-H, Chen C-T, Wang M-H, Wu C-F. A novel fault diagnosis method based-on modified neural networks for photovoltaic systems. In: International Conference in Swarm intelligence (ICSI), Beijing, China; 12-15 Jun; 2010. p. 531–9.
[55] Li X, Yang P, Ni J, Zhao J. Fault diagnostic method for PV array based on improved wavelet neural network algorithm. In: Proceedings of the world Congress on intelligent control and automation (WCICA), shenyang, China; 2-5 Mar; 2015. 1171–5.
[56] Liu G, Yu W, Zhu L. Condition classification and performance of mismatched photovoltaic arrays via a pre-filtered Elman neural network decision making tool. Sol Energy 2018;173:1011–24.
[57] Syafaruddin Karatepe E, Hiyama T. Controlling of artificial neural network for fault diagnosis of photovoltaic array. In: 2011 16th International Conference on intelligent system Applications to power systems, Hersonissos, Greece; 25–28 sept; 2011. p. 1–6.
[58] Jiang LL, Maskell DL. Automatic fault detection and diagnosis for photovoltaic systems using combined artificial neural network and analytical based methods. In: International Joint Conference on neural networks (IJCNN), Killarney, Ireland; 12–16 July 2015; 2015. p. 1–8.
[59] Mekki H, Mellit A, Salhi H. Artificial neural network-based modelling and fault detection of partial shaded photovoltaic modules. Simulat Model Pract Theor 2016;67:1–13.
[60] De Benedetti M, Leonardi F, Messina F, Santoro C, Vasilakos A. Anomaly detection and predictive maintenance for photovoltaic systems. Neurocomputing 2018;310:59–68.
[61] Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep Belief nets. Neural Comput 2006;18:1527–54.
[62] Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86:2278–324.
[63] Szegedy C, Liu Wei, Jia Yangqing, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: IEEE Conference on computer vision and pattern recognition (CVPR), Boston, MA, USA; 8-10 Jun. 2015; 2015. p. 1–9.
[64] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. San Diego, CA, USA: International Conference on Learning Representations; 7-9 May 2015.
[65] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE computer society Conference on computer vision and pattern recognition, Columbus, Ohio, USA; 24-27 June; 2014. p. 580–7.
[66] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: IEEE Conference on computer vision and pattern recognition (CVPR), Las Vegas, NV, USA; 27-30 Jun. 2016; 2016. p. 770–8.
[67] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017;60:84–90.
[68] Oktay O, Schlemper J, Folgoc L Le, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D. Attention U-Net: learning where to look for the Pancreas. Amsterdam, Netherlands: Medical Imaging with Deep Learning; 4-6 July 2018. p. 1–10.
[69] Redmon J, Farhadi A. YOLOv3: an incremental improvement. 2018.
[70] Qian X, Li J, Cao J, Wu Y, Wang W. Micro-cracks detection of solar cells surface via combining short-term and long-term deep features. Neural Network 2020; 127:132–40.
[71] Li X, Yang Q, Lou Z, Yan W. Deep learning based module defect analysis for large-scale photovoltaic farms. IEEE Trans Energy Convers 2019;34:520–9.
[72] Deitsch S, Christlein V, Berger S, Buerhop-Lutz C, Maier A, Gallwitz F, Riess C. Automatic classification of defective photovoltaic module cells in electroluminescence images. Sol Energy 2019;185:455–68.
[73] Karimi AM, Fada JS, Hossain MA, Yang S, Peshek TJ, Braid JL, French RH. Automated pipeline for photovoltaic module electroluminescence image processing and degradation feature classification. IEEE J Photovoltaics 2019;9: 1324–35.
[74] Du B, He Y, Duan J, Zhang Y. Intelligent classification of silicon photovoltaic cell defects based on eddy current thermography and convolution neural network. IEEE Trans Ind Informatics 2019. 1–1.
[75] Akram MW, Li G, Jin Y, Chen X, Zhu C, Ahmad A. Automatic detection of photovoltaic module defects in infrared images with isolated and develop-model transfer deep learning. Sol Energy 2020;198:175–86.
[76] Huerta Herraiz Á, Pliego Marugán A, García Márquez FP. Photovoltaic plant condition monitoring using thermal images analysis by convolutional neural network-based structure. Renew Energy 2020;153:334–48.
[77] Sun M, Lv S, Zhao X, Li R, Zhang W, Zhang X. Defect detection of photovoltaic modules based on convolutional neural network. In: International Conference on Machine learning and intelligent Communications, weihai, China; 6-8 July 2018. p. 122–32.
[78] Bengio Y, Boulanger-Lewandowski N, Pascanu R. Advances in optimizing recurrent networks. In: IEEE International Conference on acoustics, speech and signal processing, Vancouver, BC, Canada; 26-31 May 2013; 2013. p. 8624–8.
[79] Banda P, Barnard L. A deep learning approach to photovoltaic cell defect classification. In: Proceedings of the annual Conference of the South African Institute of computer scientists and information technologists on (SAICSIT), Port elizabeth, South Africa; 26-28 sept; 2018. p. 215–21.
[80] Wilson DR, Martinez TR. The general inefficiency of batch training for gradient descent learning. Neural Network 2003;16:1429–51.
[81] Pierdicca R, Malinverni ES, Piccinini F, Paolanti M, Felicetti A, Zingaretti P. Deep convolutional neural network for automatic detection of damaged photovoltaic cells. ISPRS - Int Arch Photogramm Remote Sens Spat Inf Sci 2018;XLII–2: 893–900.
[82] Bartler A, Mauch L, Yang B, Reuter M, Stoicescu L. Automated detection of solar cell defects with deep learning. In: 2018 26th European signal processing Conference (EUSIPCO), Rome, Italy; 3-7 sept; 2018. 2035–9.
[83] Buerhop-Lutz C, Brabec CJ, Camus C, Hauch J, Doll B, Berger S, Gallwitz F, Maier A, Deitsch S. A benchmark for visual identification of defective solar cells in electroluminescence imagery. In: 35th European photovoltaic solar energy Conference and exhibition, Brussels, Belgium; 24-28 sept; 2018. 1287–9.
[84] Dunderdale C, Brettenny W, Clohessy C, van Dyk EE. Photovoltaic defect classification through thermal infrared imaging using a machine learning approach. Prog Photovoltaics Res Appl 2020;28:177–88.
[85] Mehta S, Azad AP, Chemmengath SA, Raykar V, Kalyanaraman S. DeepSolarEye: power loss prediction and weakly supervised soiling localization via fully convolutional networks for solar panels. In: IEEE winter Conference on applications of computer vision (WACV), vol. 2018- Janua, lake tahoe, NV, USA; 12-15 Mar. 2018; 2018. p. 333–42.
[86] Ni B, Zou P, Li Q, Chen Y. Intelligent defect detection method of photovoltaic modules based on deep learning. In: Proceedings of the 2018 International

Conference on transportation & Logistics, information & communication, smart city. Paris, France: TLICSC 2018); 2018.
[87] Wei S, Li X, Ding S, Yang Q, Yan W. Hotspots Infrared detection of photovoltaic modules based on Hough line transformation and Faster-RCNN approach. In: 6th International Conference on control, decision and information technologies (CoDIT), 23-26 apr. 2019; 2019. p. 1209–14.
[88] Zhang X, Hao Y, Shangguan H, Zhang P, Wang A. Detection of surface defects on solar cells by fusing Multi-channel convolution neural networks. Infrared Phys Technol 2020;108:103334.
[89] Rahman MRU, Chen H. Defects inspection in polycrystalline solar cells electroluminescence images using deep learning. IEEE Access 2020;8:40547–58.
[90] Wang J, Zhao B, Yao X. PV abnormal shading detection based on convolutional neural network. In: Proceedings of the 32nd Chinese control and decision Conference (CCDC 2020), Hefei, China; 22-24 aug; 2020. 1580–3.
[91] Chen H, Pang Y, Hu Q, Liu K. Solar cell surface defect inspection based on multispectral convolutional neural network. J Intell Manuf 2018;31:453–68.
[92] Akram MW, Li G, Jin Y, Chen X, Zhu C, Zhao X, Khaliq A, Faheem M, Ahmad A. CNN based automatic detection of photovoltaic cell defects in electroluminescence images. Energy 2019;189.
[93] Tang W, Yang Q, Xiong K, Yan W. Deep learning based automatic defect identification of photovoltaic module using electroluminescence images. Sol Energy 2020;201:453–60.
[94] Malof JM, Collins LM, Bradbury K, Newell RG. A deep convolutional neural network and a random forest classifier for solar photovoltaic array detection in aerial imagery. In: 2016 IEEE International Conference on renewable energy research and applications (ICRERA), Birmingham, UK; 20-23 Nov; 2016. p. 650–4.
[95] Lu X, Lin P, Cheng S, Lin Y, Chen Z, Wu L, Zheng Q. Fault diagnosis for photovoltaic array based on convolutional neural network and electrical time series graph. Energy Convers Manag 2019;196:950–65.
[96] Manohar M, Koley E, Ghosh S. Enhancing the reliability of protection scheme for PV integrated microgrid by discriminating between array faults and symmetrical line faults using sparse auto encoder. IET Renew Power Gener 2019;13:308–17.
[97] Chen Z, Chen Y, Wu L, Cheng S, Lin P. Deep residual network based fault detection and diagnosis of photovoltaic arrays using current-voltage curves and ambient conditions. Energy Convers Manag 2019;198:111793.
[98] Gao W, Wai RJ. A novel fault identification method for photovoltaic array via convolutional neural network and residual gated recurrent unit. IEEE Access 2020;8:159493–510.
[99] Hopwood MW, Gunda T, Seigneur H, Walters J. Neural network-based classification of string-level IV curves from physically-induced failures of photovoltaic modules. IEEE Access 2020;8:161480–7.
[100] Rioul O, Duhamel P. Fast algorithms for discrete and continuous wavelet transforms. IEEE Trans Inf Theor 1992;38:569–86.
[101] Lu S, Sirojan T, Phung BT, Zhang D, Ambikairajah E, DA-DCGAN. An effective methodology for DC series arc fault diagnosis in photovoltaic systems. IEEE Access 2019;7:45831–40.
[102] Aziz F, Ul Haq A, Ahmad S, Mahmoud Y, Jalal M, Ali U. A novel convolutional neural network-based approach for fault classification in photovoltaic arrays. IEEE Access 2020;8:41889–904.
[103] Jones CB, Stein JS, Gonzalez S, King BH. Photovoltaic system fault detection and diagnostics using Laterally Primed Adaptive Resonance Theory neural network. In: 2015 IEEE 42nd photovoltaic specialist Conference (PVSC), new Orleans, LA, USA; 14-19 Jun; 2015. p. 1–6.
[104] Liao Z, Wang D, Tang L, Ren J, Liu Z. A heuristic diagnostic method for a PV system: triple-layered particle swarm optimization–back-propagation neural network. Energies 2017;10:226.
[105] Zhang JR, Zhang J, Lok TM, Lyu MR. A hybrid particle swarm optimization-back-propagation algorithm for feedforward neural network training. Appl Math Comput 2007;185:1026–37.
[106] Zheng YL, Lin PJ, Yu JL, Lai YF, Lin YH, Chen ZC, Wu LJ, Cheng SY, Chen GD. A novel fault diagnosis method for photovoltaic array based on BP-Adaboost strong classifier. IOP Conf Ser Earth Environ Sci 2018;188:012110.
[107] Freund Y, Iyer R, Schapire RE, Singer Y, Dietterich TG. An efficient boosting algorithm for combining preferences, vol. 4; 2003.
[108] Mohamed A H, Nassar AMN AM. New algorithm for fault diagnosis of photovoltaic energy systems. Int J Comput Appl 2015;114:26–31.
[109] Whitley D. A genetic algorithm tutorial. Stat Comput 1994;4:65–85.
[110] Chen Z, Wu L, Cheng S, Lin P, Wu Y, Lin W. Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics. Appl Energy 2017;204:912–31.
[111] Lagarias JC, Reeds JA, Wright MH, Wright PE. Convergence properties of the nelder–mead Simplex method in low dimensions. SIAM J Optim 1998;9:112–47.
[112] Wu Y, Chen Z, Wu L, Lin P, Cheng S, Lu P. An intelligent fault diagnosis approach for PV array based on SA-RBF kernel extreme learning machine. Energy Procedia 2017;105:1070–6.
[113] Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing. Science (80-) 1983;220:671–80.
[114] Bharath KVS, Haque A, Khan MA. Condition monitoring of photovoltaic systems using machine leaming techniques. In: 2018 2nd IEEE International Conference on power electronics, intelligent control and energy systems (ICPEICES), New Delhi, India; 22-24 Oct; 2018. 870–5.
[115] Heil CE, Walnut DF. Continuous and discrete wavelet transforms. SIAM Rev 1989; 31:628–66.
[116] Karmacharya IM, Gokaraju R. Fault location in ungrounded photovoltaic system using wavelets and ANN. IEEE Trans Power Deliv 2018;33:549–59.
[117] Harti A. Discrete multi-resolution analysis and generalized wavelets. Appl Numer Math 1993;12:153–92.
[118] Huang Z, Wang Z, Zhang H. Multiple open-circuit fault diagnosis based on multistate data processing and subsection fluctuation analysis for photovoltaic inverter. IEEE Trans Instrum Meas 2018;67:516–26.
[119] Zhu H, Lu L, Yao J, Dai S, Hu Y. Fault diagnosis approach for photovoltaic arrays based on unsupervised sample clustering and probabilistic neural network model. Sol Energy 2018;176:395–405.
[120] Krishnapuram R, Keller JM. A possibilistic approach to clustering. IEEE Trans Fuzzy Syst 1993;1:98–110.
[121] Kurukuru VSB, Haque A, Khan MA, Tripathy AK. fault classification for photovoltaic modules using thermography and machine learning techniques. In: 2019 International Conference on computer and information sciences (ICCIS), aljouf, Saudi arabia; 10–11 apr; 2019. p. 1–6.
[122] V BS. Grey level Co-occurrence matrices: generalisation and some new features. Int J Comput Sci Eng Inf Technol 2012;2:151–7.
[123] Lo Sciuto G, Napoli C, Capizzi G, Shikler R. Organic solar cells defects detection by means of an elliptical basis neural network and a new feature extraction technique. Optik 2019;194:163038.
[124] Klema VC, Laub AJ. The singular value decomposition: its computation and some applications. IEEE Trans Automat Contr 1980;25:164–76.
[125] Hanafy WA, Pina A, Salem SA. Machine learning approach for photovoltaic panels cleanliness detection. In: 2019 15th International computer engineering Conference (ICENCO), giza, Egypt; 29-30 Dec; 2019. p. 72–7.
[126] Balzategui J, Eciolaza L, Arana-Arexolaleiba N, Altube J, Aguerre JP, Legarda-Ereño I, Apraiz A. Semi-automatic quality inspection of solar cell based on Convolutional Neural Networks. In: 24th IEEE Conference on emerging technologies and factory automation, vol. 2019- septe, Zaragoza, Spain; 10-13 sept; 2019. p. 529–35.
[127] Tan Y, Liao K, Bai X, Deng C, Zhao Z, Zhao B. Denoising convolutional neural networks based dust accumulation status evaluation of photovoltaic panel. In: IEEE International Conference on energy Internet (ICEI), Nanjing, China; 20-24 May 2019; 2019. p. 560–6.
[128] Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. IEEE Trans Image Process 2017;26: 3142–55.
[129] Li X, Yang Q, Wang J, Chen Z, Yan W. Intelligent fault pattern recognition of aerial photovoltaic module images based on deep learning technique. In: 9th international multi-conference on complexity, informatics and cybernetics, vol. 1. Orlando, FL, USA: IMCIC; 2018. 13-16 Mar. 2018, p. 22–7.
[130] Li X, Li W, Yang Q, Yan W, Zomaya AY. Building an online defect detection system for large-scale photovoltaic plants. In: Proceedings of the 6th ACM International Conference on systems for energy-efficient Buildings, Cities, and transportation, New York, NY, USA; 13-14 Nov; 2019. p. 253–62.
[131] Qian X, Li J, Cao J, Wu Y, Wang W. Micro-cracks detection of solar cells surface via combining short-term and long-term deep features. Neural Network 2020;127: 132–40.
[132] Yan J, Zhu M, Liu H, Liu Y. Visual saliency detection via sparsity pursuit. IEEE Signal Process Lett 2010;17:739–42.
[133] Narayan S, Tagliarini G. An analysis of underfitting in MLP networks. In: 2005 IEEE International Joint Conference on neural networks (IJCNN), vol. 2, montreal, Que., Canada; 31 July-4 aug; 2005. p. 984–8.
[134] Forman SE. Using measurements to detect electrical problems in operational photovoltaic arrays. NASA STI/Recon Tech Rep N 1981;83.
[135] Dumas LN, Shumka A. Photovoltaic module reliability improvement through application testing and failure analysis. IEEE Trans Reliab 1982;R-31:228–34.
[136] Hamdy MA, Beshir ME, Elmasry SE. Reliability analysis of photovoltaic systems. Appl Energy 1989;33:253–63.
[137] Deng J, Dong W, Socher R, Li L-J, Li Kai, Fei-Fei Li. ImageNet: a large-scale hierarchical image database. In: 2009 IEEE Conference on computer vision and pattern recognition, miami, FL, USA; 20-25 Jun; 2009. p. 248–55.
[138] Dong G, Liao G, Liu H, Kuang G. A review of the autoencoder and its variants: a comparative perspective from target recognition in synthetic-aperture radar images. IEEE Geosci Remote Sens Mag 2018;6:44–68.
[139] Happel BLM, Murre JMJ. Design and evolution of modular neural network architectures. Neural Network 1994;7:985–1004.
[140] Lipton ZC, Berkowitz J, Elkan C. A critical review of recurrent neural networks for sequence learning. 2015.
[141] Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. In: 36th International Conference on Machine learning (ICML), Long Beach, CA, USA; 9-15 Jun; 2019. 10691–700.
[142] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: 30th IEEE Conference on computer vision and pattern recognition (CVPR), Honolulu, HI, USA; 21-26 July 2017. p. 2261–9.
[143] Gholami A, Kwon K, Wu B, Tai Z, Yue X, Jin P, Zhao S, Keutzer K. SqueezeNext: hardware-aware neural network design. In: IEEE/CVF Conference on computer vision and pattern recognition workshops (CVPRW), salt lake city, UT, USA; 18-22 June 2018; 2018. 1719–171909.
[144] Ma N, Zhang X, Zheng H-T, Sun J. ShuffleNet V2: practical guidelines for efficient CNN architecture design. In: European Conference on computer vision, munich, Germany; 8-14 sept; 2018. p. 122–38.
[145] Dhimish M, Holmes V, Dales M. Parallel fault detection algorithm for grid-connected photovoltaic plants. Renew Energy 2017;113:94–111.
[146] Li M, Yuan B. 2D-LDA: a statistical linear discriminant analysis for image matrix. Pattern Recogn Lett 2005;26:527–32.

[147] Domeniconi C, Papadopoulos D, Gunopulos D, Ma S. Subspace clustering of high dimensional data. SIAM Proceedings Series 2004;6:517–21.
[148] Jiang G, He H, Yan J, Xie P. Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox. IEEE Trans Ind Electron 2019;66:3196–207.
[149] Wong PK, Yang Z, Vong CM, Zhong J. Real-time fault diagnosis for gas turbine generator systems using extreme learning machine. Neurocomputing 2014;128:249–57.
[150] Ben Ali J, Fnaiech N, Saidi L, Chebel-Morello B, Fnaiech F. Application of empirical mode decomposition and artificial neural network for automatic bearing fault diagnosis based on vibration signals. Appl Acoust 2015;89:16–27.