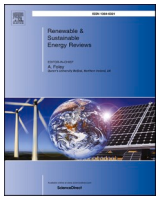
Renewable and Sustainable Energy Reviews 202 (2024) 114617

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

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



Machine learning for advanced characterisation of silicon photovoltaics: A comprehensive review of techniques and applications

Yoann Buratti [1], Gaia M.N. Javier [1,**], Zubair Abdullah-Vetter [1,*], Priya Dwivedi, Ziv Hameiri

University of New South Wales, Sydney, NSW, 2052, Australia

ARTICLE INFO

Keywords:
Photovoltaics
Silicon solar cells
Characterisation
Machine learning
Deep learning
Optimisation
Luminescence
Defects
Artificial intelligence

ABSTRACT

Accurate and efficient characterisation techniques are essential to ensure the optimal performance and reliability of photovoltaic devices, especially given the large number of silicon solar cells produced each day. To unlock valuable insights from the amount of data generated during the characterisation process, researchers have increasingly turned to different machine learning (ML) techniques. In this review, advances in ML applications for silicon photovoltaic (PV) characterisation from 2018 to 2023, including device investigation, process optimisation, and manufacturing line assessment are examined. Additionally, studies on deep learning techniques for luminescence-based measurements, such as defect classification, detection, and segmentation, which can help manufacturers identify potential reliability issues are explored. Despite the abundance of ML applications, it is emphasised that the lack of both publicly available datasets and the uniform use of ML metrics poses a significant challenge for researchers to benchmark their frameworks and achieve consistent and accurate results. In advancing ML applications in PV, future research should focus on improving model interpretability, balancing speed and accuracy, understanding computational demands, and integrating niche applications into a unified framework. Lastly, industry involvement and interdisciplinary collaboration among experts in solar energy, data science, and engineering are vital in tailoring ML solutions and enhancing innovation in addressing various challenges in the PV field.

Abbreviations

AB	Adaptive boosting
AE	Autoencoder
AI	Artificial intelligence
Al-BSF	Aluminium back surface field
AlO_x	Aluminium oxide
a-Si	Amorphous silicon
B	Boron
BAF	Bi-directional attention feature pyramid network
CAN	Complementary attention network
CIGS	Copper indium gallium selenide
CNN	Convolutional neural network
CV	Cross-validation
Cz-Si	Czochralski silicon
DBN	Deep belief network
DEA	Differential evolution algorithm
EL	Electroluminescence
EQE	External quantum efficiency

(*continued on next column*)

(*continued*)

ERDCF	Efficient and refined deep convolutional features
FDTD	Finite difference time domain
Fe	Iron
FEM	Finite element method
FF	Fill factor
FN	False positive
FP	False negative
FPN	Feature pyramid network
GA	Genetic algorithm
GAB	Global attention block
GAN	Generative adversarial network
GAR	Guided anchoring region
GR	Gaussian regression
HJT	Heterojunction
IoU	Intersection over union
IPCC	Intergovernmental panel on climate change
IRT	Infrared thermography
ITRPV	International roadmap for photovoltaics

(*continued on next page*)

* Corresponding author.
** Corresponding author.
E-mail addresses: g.javier@unsw.edu.au (G.M.N. Javier), z.abdullahvetter@unsw.edu.au (Z. Abdullah-Vetter), z.hameiri@unsw.edu.au (Z. Hameiri).
[1] Equal contribution.

https://doi.org/10.1016/j.rser.2024.114617
Received 17 September 2023; Received in revised form 17 May 2024; Accepted 1 June 2024
Available online 1 July 2024

(*continued*)

IQE	Internal quantum efficiency
I–V	Current-voltage
k-NN	k-nearest neighbours
LDA	Linear discriminant analysis
LSTM	Long short-term memory
MAE	Mean absolute error
mAP	Mean average precision
mc-Si	Multicrystalline silicon
ML	Machine learning
mono-Si	Monocrystalline silicon
NMA	Nelder-Mead algorithm
NN	Neural network
OES	Optical emissions spectroscopy
PCA	Principal component analysis
PECVD	Plasma-enhanced chemical vapor deposition
PERC	Passivated emitter and rear contact
PL	Photoluminescence
$POCl_3$	Phosphorus oxychloride
PSALHE-EA	Parallel self-adaptive low-high evaluation evolutionary algorithm
PSO	Particle swarm optimisation
PV	Photovoltaic
R–CNN	Region-proposal convolutional neural network
RF	Random forest
RMSE	Root mean square error
RSA	Random search algorithm
SAA	Simulated annealing algorithm
S-ALD	Spatial atomic layer deposition
SE	Squeeze and excitation blocks
SEM	Structural equation modelling
Si	Silicon
SiN_x	Silicon nitride
SiO_x	Silicon oxide
SRH	Shockley Read Hall
SSD	Single-shot detection
SSIM	Structure similarity index
SVM	Support vector machine
TN	True negative
TP	True positive
UVF	Ultra-violet fluorescence
YOLO	You only look once

Symbols	
$b_{i,c}$	Binary indicator of element i in class c
F_1	Harmonic mean of precision and recall
I_{sc}	Short-circuit current [A]
J_{sc}	Short-circuit current density [A/cm^2]
M	Number of classes
N	Dataset size
$p_{i,c}$	Predicted probability of element i in class c
P_{mp}	Maximum power point power [W]
R^2	Coefficient of determination
RE_{max}	Maximum relative error
R_s	Series resistance [Ω]
R_{sh}	Shunt resistance [Ω]
V_{oc}	Open-circuit voltage [V]
y_i	i-th value
y_{mean}	Mean value
y^{pred}	Predicted value
y^{true}	Actual value
σ_n	Electron capture cross-section [cm^2]
σ_p	Hole capture cross-section [cm^2]
E_t	Defect energy level [eV]
k	Capture cross-section ratio

1. Introduction

Nearly 75 % of greenhouse gases emitted into the earth's atmosphere are due to human activities, primarily from the burning of fossil fuels [1]. This has caused a significant impact on the climate, and it is vital to move towards a carbon-free society, starting with the energy sector. Solar energy, such as photovoltaics (PV), is a natural solution to address the rise in global temperatures due to the abundance of solar energy reaching the planet's surface [2]. The Intergovernmental Panel on Climate Change (IPCC) recommends that one-third of the planet's energy should come from PV by 2030 [1]. Additionally, the production and installation costs of PV have significantly decreased over the past decade, making it the cheapest source of electricity [3]. To accelerate the transition toward solar-based electricity generation, the international technology roadmap for photovoltaics (ITRPV) has identified three key strategies: cost optimisation across the PV production value chain, improved efficiency and reliability of PV devices, and specialised PV modules for different market segments [4].

To address the first two ITRPV's key strategies, i.e., to ensure continuous optimisation and improvement in the efficiency and reliability of PV devices, high-quality characterisation is necessary. PV production lines employ various characterisation techniques such as current-voltage (I–V) measurements and luminescence imaging. I–V measurements provide key electrical parameters while luminescence images, either electroluminescence (EL) [5] or photoluminescence (PL) [6], provide spatial information such as local defects. These "inline" characterisation techniques are crucial to ensuring efficient production lines and optimising processes to produce reliable solar cells. Moreover, characterisation can also be done "offline", to provide more insight into efficiency losses or material properties. Finally, characterisation techniques such as EL, PL, infrared thermography (IRT) [7], or ultra-violet fluorescence imaging (UVF) [8] are also applied to PV modules installed in the field to assess their performance.

The PV market is currently dominated by crystalline silicon (Si) technologies, which account for 95 % of the market share [9]. Within the Si-based technologies, mono-crystalline Si (mono-Si) wafers dominate the market [10]. Passivated emitter and rear contact (PERC) cells have dominated the PV market in the last few years [9,11]. However, tunnel oxide passivated contact (TOPCon) and silicon heterojunction (SHJ) solar cells have recently gained a significant market share [11]. While other technologies such as thin films or perovskites [10] are also present in the market, the ITRPV [9] predicts that Si-based technology will remain the market leader over the next decade.

The development of artificial intelligence (AI) has brought a new set of tools that has revolutionised various industries, including healthcare [12], engineering [13], and education [14]. With the advent of machine learning (ML), the use of AI has the potential to significantly impact the PV industry. By accelerating the development of breakthrough innovations, and improving current solutions, AI can be a game-changer for the PV industry. In particular, AI could play a crucial role in PV characterisation, streamlining methods, automating analysis, and optimising production lines. By leveraging AI, the PV industry can improve efficiency, reduce costs, and achieve higher levels of performance and reliability. This could result in scenarios where even slight improvements in material efficiency, identified through ML-powered analysis, can lead to substantial increases in energy output over the lifetime of PV technologies, thereby enhancing their economic value ($/W).

Previously published reviews on ML in PV are presented in Table 1, with a focus on their application, support level, type of data used, and techniques covered. For techniques, classical regression and classification algorithms include comparatively basic algorithms such as random forests (RF) and support vector machines (SVM). Optimisation algorithms include the genetic algorithm (GA), particle swarm optimisation (PSO), and the like. Deep learning involves more complex architectures like convolutional neural networks (CNN).

The most comprehensive review by Mellit and Kalogirou [15] covers ML and optimisation algorithms across multiple applications and data types, with recommendations mainly concentrated on system-level applications. Among the ML applications, power forecasting of modules and PV systems is the most prominent [15,17,19–22,24,25,30], followed by fault identification and prediction in PV systems [24,30,18,23,27]. Optimisation algorithms were primarily reviewed for data-driven modelling and system design [15,17,21,30]. Recent reviews have shifted their focus to ML and deep learning applications for defect detection and classification from image-based data [23,26–29]. It is worth noting that previous reviews place a heavy focus on system-level applications,

Table 1
A comparison of recent reviews for various ML applications in PV.

Reference	Year	Focus					Application			Type of data			AI techniques covered		
		Power forecasting	Data-driven modelling	System design and sizing	Fault identification and prediction	Defect detection and classification	Cell level	Module level	System level	Time series data	Numerical data	Image-based data	Classical regression and classification algorithms	Optimisation algorithms	Deep learning
Mellit and Kalogirou [15]	2008	X	X	X				X	X	X	X	X	X	X	
Yang et al. [16]	2017	X							X	X	X	X	X	X	X
Youssef et al. [17]	2017	X	X	X			X		X		X		X	X	
AbdulMawjood et al. [18]	2018				X			X	X	X		X	X		
Ahmed et al. [19]	2020	X							X	X			X		X
Rajagukguk et al. [20]	2020	X							X	X					X
Garud et al. [21]	2021	X	X						X	X	X		X	X	
Ridha et al. [22]	2021	X		X					X	X			X	X	
Li et al. [23]	2021				X	X	X	X	X	X		X	X		X
Navid et al. [24]	2021	X			X				X	X		X	X		
Mellit and Kalogirou [25]	2021				X				X	X	X		X		X
Al-Mashhadani et al. [26]	2021					X		X				X	X		X
Berghout et al. [27]	2021				X	X		X	X	X	X	X			X
Hong and Pula [28]	2022				X	X			X	X	X	X	X		
Akram et al. [29]	2022					X		X				X	X		X
Gaviria et al. [30]	2022	X		X	X				X	X	X		X		
This review	2024		X			X	X	X			X	X	X	X	X

neglecting cell-level applications. The result then is that there has not been a review that combines ML, deep learning, and optimisation techniques for PV characterisation at the cell level.

This review aims to bridge the gap in the literature by providing a comprehensive analysis of recent advancements in using ML for PV characterisation between 2018 and 2023, with a focus on Si-based technology and manufacturing-related characterisation. Field characterisation after installation is excluded from this review. Additional background information on ML concepts has been provided in the supporting documentation. This includes a copy of the abbreviations and symbols list, and an overview of ML concepts, including various algorithms, training procedures, and scoring metrics. This review explores how ML is used to optimise processes, model and analyse devices, and assess manufacturing lines. In addition, this review conducts an in-depth analysis of deep learning applications in luminescence imaging. Studies are compared using identical metrics, where feasible, enabling the extraction of general trends. Finally, this review provides recommendations for future research in the field.

2. Classical ML applications

Big data analytics has emerged as a critical tool for reducing costs and improving decision-making in solar cell production lines. Various measurement techniques and ML algorithms can be utilised to provide valuable insights for PV manufacturers and researchers alike. This section provides an overview of prior research in the field of classical ML applications on process optimisation, device investigation, and manufacturing line assessment.

2.1. Process optimisation

Optimising the processes involved in Si solar cell manufacturing is crucial for improving the performance of mass-produced cells [31]. In recent years, ML regression models and optimisation approaches have been increasingly utilised for this purpose [32,33].

Early studies on ML-assisted process optimisation have been done by Kayabasi et al. [34,35] In their first study [34], they determined the optimum parameters of a diamond wire saw-cutting process for Si ingots. Twenty-eight experiments were conducted to obtain the surface roughness ('target output') as a function of different combinations of spool speed, z-axis speed, and coolant oil ('input parameters'). These data were used to train a neural network (NN) which resulted in a 1.43 % mean absolute error (MAE) in predicting the surface roughness. A parameter sweep was done to select the best configuration of input parameters that would output the lowest surface roughness. However, the optimum parameters obtained were not validated. A similar approach was used in their second study [35] wherein they tried to optimise nano-etching parameters to obtain the lowest surface reflectance of Si wafers. They conducted 32 experiments with varying masking and etching durations of n-type mono-crystalline Si wafers. The masking duration and etching duration were both used as inputs to the NN, while the surface reflectance was used as a target output. The trained model resulted in an MAE of 1.01 %. The optimum process parameters were determined through a parameter sweep and results were validated through experiments.

These studies were followed by Ozturk et al. [36,37]. Their work focused on the optimisation of the lapping or refining process of Si wafers. In their first study [36], they conducted 218 lapping operations to generate the dataset and trained their NN. Input parameters were rotation speed, lapping duration, and lapping pressure, while the output parameter was the surface roughness. A parameter sweep was performed using the trained network to find the input parameters that would minimise the surface roughness. Lastly, to validate their results, a scanning electron microscope was used to investigate the surface roughness of the sample produced from the optimum parameters. In their second study [37], they did fewer experiments where only the rotation speed and lapping duration were varied. They replaced the NN with several nonlinear regression models that resulted in a maximum coefficient of determination (R^2) of 0.997. They also improved their optimisation approach by implementing the differential evolution algorithm (DEA), Nelder-Mead algorithm (NMA), random search algorithm (RSA), and simulated annealing algorithm (SAA). RSA and DEA had the best performance among the four, however, the results were not validated through experiments or numerical simulations. This was the first study on stochastic optimisation of surface roughness of Si wafers.

An advantage of optimisation algorithms is that there is no limit on the number of parameters to be optimised. Buratti et al. [32] demonstrated this capability through their work on optimising the recipe for an aluminium back surface field (Al-BSF) production line. For their initial data, they simulated a virtual production line using Python/PC1D [38] Monte Carlo simulations [39]. Their simulations involve ten processing steps and 47 process parameters resulting in around 400,000 simulated cells. For regression, different ML algorithms were employed: support vector machine (SVM), adaptive boosting (AB), random forest (RF), and NN. Inputs were the 47 process parameters, and the target output was the efficiency of the cell. It was seen that the NN's predictions were the most accurate with a $R^2 = 0.996$ and root mean square error (RMSE) = 0.019 %. This trained model was then used as the objective function of a genetic algorithm (GA) optimiser. They observed that the optimum recipe increased the absolute efficiency by 2 %.

Qi et al. [40] used a combination of NN and GA to optimise parameters in a Czochralski Si (Cz-Si) crystal growth process. They used a global heat and mass transfer model through in-house software to generate 50 samples which they verified via experimentation. An NN was trained to replace the numerical transfer model: the inputs were the crystal rotation rate and crucible rotation rate, while the outputs were the average oxygen concentration and deflection of the interface. The max relative error (RE_{max}) was 5.27 % and 5.19 % for the two outputs. Since there were two outputs, a double-objective optimisation approach was used, and the cost function was the equally weighted sum of the two NNs. Optimum parameters were further validated with the global heat and mass transfer model. Reductions of 19.92 % and 4.43 % were observed in the deflection of the interface and the average oxygen concentration, respectively.

Rachdi and Hoffman [41] used a different approach for optimising the plasma-enhanced vapor deposition (PECVD) of silicon nitride (SiN_x). Optical emissions spectroscopy (OES) measurements were used to predict SiN_x layer properties (thickness, refractive index, bond densities) through linear regression fitting. The regression model had a RE_{max} of 10.3 % for thickness, 1.7 % for refractive index, and 35.8 % for bond densities. They extended their model by correlating OES measurements with process parameters. They then implemented a parameter sweep to obtain ideal process parameters. They discovered that optimum layer properties can be achieved by increasing the power and decreasing the amount of gas flow.

Lastly, Wagner-Mohnsen *et al.* [33] focused on optimising a phosphorus oxychloride ($POCl_3$) diffusion process. Sentaurus was used to generate 2,400 sets of data. A Gaussian regression (GR) model was then trained to predict the solar cell efficiency using several input parameters: time, temperature, etch depth, and finger pitch. They implemented a five-fold cross-validation method and the resulting RMSE was 0.085 %. To determine the optimum process parameters, the GR model was used as an objective function of a GA. After approximately 150 iterations, the final input parameters improved the solar cell efficiency to around 23.4 %.

ML can be used to optimise different processes in manufacturing solar cells. In terms of data, early studies have relatively small experimental datasets (<100), but the accuracy of their regression models was acceptable. Conversely, recent studies like Wagner-Mohnsen *et al.* [33] made use of simulations that allowed the generation of large datasets (>1,000). However, a common metric for assessing the accuracy of the regression models has not been established.

To date, most optimisation approaches have used a parameter sweep.

Table 2
Studies on ML-assisted process optimisation (*calculated from data given).

Reference	Year	Focus	Regression model	Accuracy	Optimisation approach
Kayabasi et al. [34]	2017	Diamond wire saw cutting	NN	MAE = 1.43 %* R^2 = 0.993*	Parameter sweep
Kayabasi et al. [35]	2018	Wet chemical etching	NN	MAE = 1.01 %* R^2 = 0.999	Parameter sweep
Ozturk et al. [36]	2018	Wafer lapping process	NN	MAE = 2.72 %* R^2 = 0.998*	Parameter sweep
Ozturk et al. [37]	2018	Wafer lapping process	Non-linear regression	R^2 = 0.997	DEA, NMA, RSA, SAA
Buratti et al. [32]	2020	Entire line fabrication	ANN, RF, AB, SV	RMSE = 0.019 % R^2 = 0.996	GA
Qi et al. [40]	2020	Cz-Si crystal growth	NN	RE_{max} = 5.27 %, 5.19 % R^2 = 0.974*	GA
Rachdi & Hofmann [41]	2021	PECVD	Linear/Non-linear regression	RE_{max} = 10.3 %, 1.7 %, 35.8 %	Parameter sweep
Wagner-Mohnsen et al. [33]	2021	$POCl_3$ diffusion	GR	RMSE = 0.085 %	GA

This is acceptable if only a few input parameters are being used for the optimisation procedure, however, when more input parameters are being investigated, stochastic optimisation algorithms like GA are preferred to consider various interactions between the input parameters [42]. Most of the studies did not validate their results with experiments or simulations. It is recommended to conduct this validation to enable further investigation and strengthen the findings. Additionally, publishing the code and/or datasets is uncommon. Making these resources publicly available aids other researchers in further developing and validating their work. A summary of research on ML-assisted process optimisation is presented in Table 2.

2.2. Device investigation

Solar cells have a complex structure with several factors contributing to their performance. Device investigation allows researchers and manufacturers to understand different loss mechanisms and further optimise solar cell performance.

2.2.1. Simulation-based applications

ML has been used to model the physics of different aspects of a solar cell. The following studies feature a wide range of objectives from replicating computationally expensive solar cell simulation models to simplifying the optimisation process of solar cell design with the assistance of ML. While these studies are quite varied, they all show that ML is applicable in different aspects of solar cell characterisation and design optimisation pipelines by replacing or enhancing existing physics models.

Aiello et al. [43] applied a parallel self-adaptive low-high evaluation evolutionary algorithm (PSALHE-EA) to optimise the structure of a thin-film Si solar cell. Parameters that were investigated were the radius and location of silica and silver nanoparticles on top of the cell. They developed an objective function based on finite element method (FEM) simulations for light scattering to maximise the efficiency and at the same time, minimise the number of silver nanoparticles. Through the PSALHE-EA, the proposed optimum designs for varying cell thicknesses yielded suitable efficiency values with minimal silver consumption.

Kaya and Hajimirza [44] aimed to replace finite difference time domain (FDTD) simulations with ML. Usually, FDTD simulations are used for wide bandwidth studies, which can be computationally expensive. Studying the optimal design for thin film amorphous Si (a-Si) solar cells, they used a trained NN to predict the external quantum efficiency (EQE) of the simulated cells as a surrogate model in place of FDTD. The total training dataset consisted of 1,000 FDTD simulations, and cross-validation with four folds was used to train the NN to replace the FDTD. The design of the a-Si solar cell was optimised using SAA and the quasi-Newton method. While it is not clear if there was a separate test set after identifying the optimum NN structure, the trained model allowed the optimisation procedure to be 20 times faster than using FDTD to simulate the EQE of the solar cell design. Furthermore, the ML-optimised cell obtained a 25 % improvement compared to a randomised cell design.

Hamedi et al. [45] developed an ML model to predict the spectral absorption of nanowires used in solar cells. They replaced FDTD simulations by training an NN on data with the pitch and radius as input and the absorption spectrum as an output (200 data points with values of the target vector across 300–500 nm). The optimised NN achieved a minimum RMSE of 0.017 nm on the simulated data and provided the output predictions faster than the computationally expensive FDTD modelling approach.

Buratti et al. [46] proposed an approach to invert the Shockley–Read–Hall (SRH) [47,48] equation. They fed simulated lifetime curves into ML models to predict the band energy of the defect (E_t) and the ratio (k) of electron and hole capture cross-sections (σ_n, σ_p). Over a million lifetime curves were generated by varying E_t, k, temperature, and excess carrier concentration in the SRH equation. State-of-the-art regression models such as RF, adaptive boosting (AB), NN, and SVM models were trained on the simulated lifetime curves and an average R^2 of 0.99 was achieved. Their results were also validated on experimental data.

A similar methodology was proposed by Abdullah-Vetter et al. [49] to automate the analysis of internal quantum efficiency (IQE) measurements. IQE curves of Si solar cells were simulated using an analytical model [50] and a series of RF models were trained to predict six key parameters from the curves. These parameters are correlated to the performance of the emitter (collection efficiency [IQE_0] and width [w_e]), bulk region (surface recombination velocity [SRV_B] and effective diffusion length [L_{eff}]), and rear optics (rear internal reflectance [R_B] and how diffused the light is after internal reflection [D_B]) of a Si solar cell. They also used the predicted parameters to decouple the parasitic absorption of the anti-reflection coating. RMSE values were reported for each of the six parameters (Table 3) and their model demonstrated low error values.

Zhu et al. [51] modelled the effect of front and rear passivating film layers on PERC solar cells using ML. They studied the passivating effects of silicon oxide (SiO_x) and SiN_x layers on both the front and rear of thin (100 μm) PERC solar cells, with the assistance of NNs. The investigated film layers were a $SiO_x/SiN_x/SiN_x$ front structure and SiO_x/SiN_x rear structure. The thickness of the five layers was the input to the NN with cell efficiency as the target output. From a training dataset of 139 simulated cells, the NN learned the correlation between the thicknesses of the thin film layers and the efficiency of the resulting cell. In a small test set of 17 cells, the NN achieved an RMSE of 0.22 % and R^2 of 0.984. They also showed that the two front SiN_x layers (input features) had the

Table 3
Studies on simulation-based applications of ML-assisted device investigation.

Reference	Year	Focus	Physics model	ML model	Main results
Aiello et al. [43]	2017	Optimise the sizing and positioning of silica and silver top cell nanoparticles	FEM simulations on light scattering	PSALHE-EA	Improved efficiency with minimal silver usage
Kaya and Hajimirza [44]	2018	Optimise the cell geometry of a-Si cells	FDTD simulations on optical absorptivity	NN, SAA, quasi-Newton	Improved EQE by 25 %–50 % and ×20 modelling speed
Hamedi et al. [45]	2019	Replace FDTD simulations when modelling nanowire solar cells	FDTD simulations on electromagnetic absorption	NN	RMSE = 0.017 nm
Buratti et al. [46]	2020	SRH equation	Invert the SRH equation to extract defect parameters	RF, AB, NN, SVM	$R^2 = 0.99$
Zhu et al. [51]	2022	Predicting efficiency from the front and rear film design	Quokka simulations	NN	RMSE = 0.22 % $R^2 = 0.984$
Abdullah-Vetter [49]	2023	IQE measurement analysis of Si solar cells	Analytical model [50] of IQE to fit IQE measurements	RF	RMSE $IQE_0 = 0.00481$ $w_e = 12.1$ nm $SRV_B = 12.2$ cm/s $L_{eff} = 0.192$ cm $R_B = 0.0192$ $D_B = 0.0383$

largest effect on the NN's efficiency predictions.

In summary, ML can be used to replace or enhance existing physical models to design or investigate PV devices. ML-powered design optimisation is a common application in this area, and it seems to be more prominent in organic PV and other more complex solar cell structures and designs. While the ML objectives described in this section are quite varied, NNs were the preferred ML algorithm. A summary of the research discussed in this section is presented in Table 3.

2.2.2. *Data-based applications*

Characterisation techniques are used to analyse solar cells and infer their underlying material properties. Over the last few years, different studies have built ML models using experimental measurements, instead of analytical models. To facilitate simpler and more efficient analyses, the resulting ML models have been employed for regression and assisted inference.

Kumar and Maheshwari [52] demonstrated the capability of using RF and gradient boosting classifiers to identify defects normally found in EL images using I–V measurements. The total dataset consisted of 6,051 module I–V and EL measurements of fielded modules from different manufacturers. The study involved extracting features such as typical I–V parameters as well as statistical features such as the skewness, kurtosis, and entropy of the current and voltage of the I–V curves. These features were input into the ML models to categorise the EL images into different defect classes. It was shown that the RF and gradient boosting classifiers achieved an F_1-score of 91.2 % and 97 %, respectively. While this study draws similarities to the studies discussed in Section 3, it provides a novel method to classify the defects found in EL images without using any EL measurements, hence, showing that the ML models can correlate the effect of these defects with I–V measurements.

Grau-Luque et al. [53] predicted the thickness of aluminium oxide (AlO_x) deposition by spatial atomic layer deposition (S-ALD) on different substrates using unsupervised ML. The study focused on proposing a new characterisation method for S-ALD on flexible PV modules, which involved 30 × 30 points of normal reflectance measurement profiles. The samples consisted of AlO_x layers deposited on copper indium gallium selenide (CIGS) and Si cells as well as backsheet material with distinct thicknesses. The resulting measurement points were converted to a control parameter via a subtraction from the reference measurement points collected before the deposition. Principal component analysis (PCA) and linear discriminant analysis (LDA) were used to categorise the individual control parameter measurement points into separate classes associated with the thickness of the AlO_x layers of the samples. It was found that different illumination sources and different reflectance measurement profiles lead to differing classification results, for example, 98 % accuracy with a halogen lamp source and 72 % accuracy with an LED source.

Kato et al. [54] used a long short-term memory (LSTM) neural network for estimating crystal orientation of grains on multi-crystalline Si (mc-Si). By analysing variations in light reflection patterns from optical images, they were able to predict crystal orientations with most errors below 10°.

The analysis of measurements from PV devices is often used to identify key performance or key parameters that define the investigated devices. A common approach was to collect large datasets and then use ML to infer the relationships among different parameters. It is also evident that the number of ML applications utilising real data is significantly lower compared to simulations. This highlights the inherent challenges and complexities involved in data collection as opposed to running simulations. Table 4 summarises the studies discussed in this section.

Within the context of device investigation, trends involving transformers and large language models (LLMs) are also noteworthy. In the PV field, LLMs have been applied to tasks such as text mining and

Table 4
Studies on data-driven applications of ML-assisted device investigation.

Reference	Year	Focus	Experimental data	ML model	Main results
Kumar and Maheshwari [52]	2021	Classifying EL defects from IV measurements	I–V measurements	RF and gradient boosting	F_1-score = 97 %
Grau-Luque et al. [53]	2021	Predicting the AlO_x layer thickness	Normal reflectance	PCA-LDA	98 % accuracy with halogen illumination, 72 % with LED
Kato et al. [54]	2022	Estimating crystal orientation of grains on mc-Si	Optical images	LSTM	Majority of errors <10°

material discovery [55–57]. While these topics are promising, they are beyond the scope of this review.

2.2.3. Hybrid applications

ML models become more robust as more data are fed to them. Several studies have built and validated ML models through a combination of analytical models and experimental measurements.

Ma et al. [58] investigated the ability of linear regression to extract I–V parameters automatically and accurately from light I–V measurements of PV modules under different bypass diode conditions. Upon activation of the bypass diodes of a PV module, the I–V curve shows characteristic steps in the current output. I–V curves were simulated using the one-diode model, splitting the steps in the curves and treating them as separate I–V curves. Linear regression upon a "five-point" moving window was applied on the different steps of the curve and the short-circuit current (I_{sc}), open-circuit voltage (V_{oc}), series resistance (R_s), shunt resistance (R_{sh}), power at maximum power point (P_{mp}), and fill factor (FF) were reported for each step of the I–V curves. The proposed method achieved MAEs of 2 % for I_{sc}, 0.5 % for V_{oc}, 21.1 % for R_s, 24.65 % for R_{sh}, 3.43 % for P_{mp}, and 2.9 % for FF from 2.2 million I–V curves of PV modules, including measurements reported by Fraunhofer-ISE laboratory [59].

Wagner-Mohnsen and Altermatt [60] used a combination of numerical device modelling (Sentaurus), ML, and statistics to understand how different material properties influence the performance of industrial PERC cells. They initially simulated 400 PERC cells through Sentaurus. In each simulation, they varied the material properties: wafer resistivity, emitter doping, front recombination velocity, SRH lifetime, dimensions of the front metal fingers, BSF depth, and rear passivation quality. With these, they extracted the corresponding I–V parameters from the simulated devices. GR, RF, and SVM were trained to replicate the Sentaurus simulations. The GR model yielded the most accurate results in terms of RMSE with values for the short-circuit current density (J_{sc}), V_{oc}, FF, and efficiency of 0.026 mA/cm^2, 0.372 mV, 0.017 %, and 0.018 %, respectively. They matched the GR model results with 15,722 actual PERC measurements to trace back the material properties of these cells. Through their approach, they were able to determine and compare the material properties of cells from the lowest efficiency bin to the highest efficiency bin.

Olikh et al. [61] demonstrated the prediction of iron (Fe) impurity characteristics. The concentration of Fe in Si cells was predicted from the ideality factor extracted from simulated dark I–V curves. The authors simulated dark I–V curves in the computational software, SCAPS-1D, by sweeping temperature, bulk doping boron (B) concentration, Fe concentration, and base thickness and fit the ideality factor using the one diode model. It was assumed that the change in the extracted ideality factor was only associated with Fe and Fe–B impurities in the bulk of the solar cell. The ideality factor was used as input to train the NN to predict the Fe concentration. They achieved a RMSE = 0.54 and R^2 = 0.813. Furthermore, the ideality factor caused by only Fe can be decoupled by two I–V measurements with the second under intense illumination to dissociate the Fe–B pairs. This procedure was done to obtain the decoupled ideality factor of Fe and used as an extra feature to re-train the ANN. The resulting model achieved better results of RMSE = 0.138 and R^2 = 0.948, however, it requires the users to conduct extra measurements.

Most hybrid-powered applications in this field are focused on parameter extraction that satisfies both physics models and experimental measurements. I–V measurements are one of the standard characterisation methods for solar cells, thus they were used in all the presented studies. The access to simulated data enhanced the performance of ML algorithms due to an increase in the training data, especially if the model used for the data generation is accurate and representative of experimental data. Furthermore, simulated data can be generated with or without noise to train the ML to be robust to measurement noise. ML algorithms in this area are becoming more sophisticated as they transition from simple polynomial regression to more complex NN. A summary of the studies is presented in Table 5.

2.3. Manufacturing line assessment

Quality control ensures that manufactured solar cells comply with industrial standards. In PV manufacturing lines, non-compliant cells are rejected while compliant cells are binned according to their performance [62].

Wasmer et al. [63] studied the impact of material and process variations on the efficiency distribution of mc-Si PERC cells. They matched efficiencies and corresponding process parameters of 800 fabricated cells through Sentaurus simulations. A GR model was trained using simulated data and then the model was used to generate more data. With a tenfold cross-validation CV, the R^2 and MAE of the trained model were 0.998 and 0.022 %, respectively. The generated data was investigated using a variance-based sensitivity analysis where material- and process-related parameters were ranked according to their impact on the efficiency variance. Through their approach, they discovered that 68 % of the efficiency variance is material-related while 22 % is process-related, explaining 90 % of the overall variance.

Klöter [64] proposed another approach to identify sources of variance in solar cell performance. He analysed 200,000 cell measurements from a production line. For each cell, 56 parameters (e.g., inline measurements and wafer position in batch) were investigated. SVM models with a radial basis function kernel were trained on these parameters and

Table 5
Studies using hybrid applications of ML-assisted device investigation.

Reference	Year	Focus	Data source	ML model	Main results
Ma et al. [58]	2019	Extracting I–V parameters of PV modules	One-diode model simulations; I–V measurements	Moving window linear regression	MAE of: I_{sc}: 2 % V_{oc}: 0.5 % R_s: 21.1 % R_{sh}: 24.65 % P_{mp}: 3.43 % FF: 2.90 %
Wagner-Mohnsen and Altermatt [60]	2020	Inferring material properties of PERC cells	Sentaurus simulations; I–V measurements	RF, SVM, GR	RMSE of: J_{sc}: 0.026 mA/cm^2 V_{oc}: 0.372 mV FF: 0.017 % Efficiency: 0.018 %
Olikh et al. [61]	2022	Extracting iron concentration from dark I–V ideality factor	SCAPS-1D simulations; I–V measurements	Custom deep NN	RMSE = 0.138 R^2 = 0.948

Table 6
Studies aiming to develop ML-assisted manufacturing line assessment.

Reference	Year	Focus	ML model	Main results
Wasmer et al. [63]	2017	Identifying sources of efficiency variance	GR, variance-based sensitivity analysis	90 % of the measured variance in efficiency explained
Klöter [64]	2018	Identifying sources of efficiency variance	SVM	High accuracy in model prediction, fast identification of variance sources
Evans and Boreland [65]	2018	Identifying sources of efficiency variance	SEM, commonality analysis	Detailed proportion of different sources of variance, flagged anomalies in a production line
Arena et al. [66]	2021	Detecting anomalies during PV manufacturing	Monte-Carlo, PCA	Estimated 1–2 % downtime reduction
Pei et al. [67]	2021	Quality monitoring and control of series solar cell production line	Factor analysis, SVM, PSO, etc.	Detailed insights on quality loss mechanisms
Wasmer et al. [68]	2021	Identifying sources of efficiency variance	Decision trees, Hierarchical clustering	Top 20 sources of variance were identified

their corresponding electrical parameters. Though no accuracy metrics were reported, a histogram of the difference between the measured and predicted V_{oc} was presented. It was shown that 90 % of the predicted values have errors of less than 0.9 mV. The trained weights of each input parameter from the SVM model were used to rank the influence of each parameter on the variance in solar cell performance. His results show that the highest impact is attributed to wafer material parameters.

Evans and Boreland [65] applied structural equation modelling (SEM) and commonality analysis on standard manufacturing line measurements to understand the relationships between parameters (e.g., grid finger width, recombination, device layer resistivity, and electrical performance) and their impact on the variability in solar cell performance. The authors presented four case studies: (1) investigation of the grid finger width, (2) impact of series resistance and recombination on the efficiency variance, (3) variance analysis with mid-line data, and (4) evaluation of the resistivity device layers and contact grid resistance. The presented method provided insights into the impact of the above parameters on the variance in the line. The fourth case study was applied in a production environment and aided manufacturers to identify inconsistent cells in their line.

Arena et al. [66] implemented Monte-Carlo and PCA to detect anomalies in a PV manufacturing line. Monte-Carlo simulations were first used to pre-process sensor measurements from different stations in the line. A PCA-based anomaly detection model was then trained using pre-processed data during normal operation. In addition, a key performance indicator was assigned to each sensor. Their approach was tested in different scenarios and demonstrated the ability to identify anomalies in unseen data. Assuming that their approach was implemented on the whole manufacturing line, they estimated a downtime reduction of 1–2 %.

Pei et al. [67] applied a combination of factor analysis, SVM, and particle swarm optimisation (PSO) to investigate the different loss parameters and their contribution to the variance in machining quality. Results in the optimisation space were looped back to the physical space, enabling optimisation and control in the actual manufacturing line.

Wasmer et al. [68] also implemented ML methods to decompose sources of efficiency variance of p-type Cz-Si wafers. The authors analysed about half a million cells with around 300 measurements (e.g., anti-reflection coating thickness, emitter sheet resistance) per cell. An ensemble of decision trees was trained on this data ($R^2 = 0.84$). Since features can be highly correlated with each other, hierarchical clustering was implemented to group similar features. Through their approach, the top 20 features/clusters that have an impact on efficiency were identified. Further investigation included studying the temporal impact of each of these features on cell efficiency.

ML-assisted manufacturing line assessment studies span various applications and datasets. The aforementioned studies applied advanced statistics and ML techniques to achieve their objectives. Most of the research works are focused on identifying different sources of efficiency variance where no accuracy metric has been defined. Other studies were focused on anomaly detection in manufacturing lines where downtime reduction [66] has been used to evaluate the models developed. Large datasets are important elements in these studies. These data are often private but some researchers started publishing their data [69], making it accessible to more researchers in the field. Future studies should include temporal analysis of ML applications in production lines as the line may deviate from the provided training data and continuous transfer learning can be utilised to overcome such an issue. Real-time monitoring also plays a crucial role in this context, and advancements like edge computing are promising for facilitating real-time data processing. A summary of the studies is presented in Table 6.

3. Deep learning applications

Deep learning has emerged as a popular approach for image analysis [70]. The approach has been heavily used in the solar cell production industry, particularly for advanced analysis of PV luminescence images

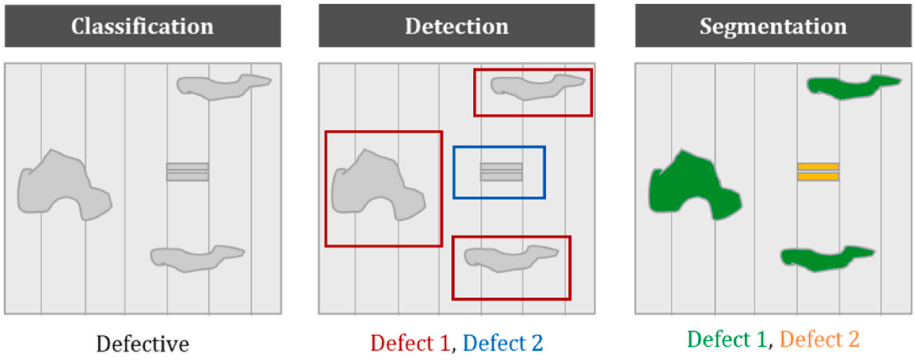


Fig. 1. Difference between defect classification, detection, and segmentation.

[71–73]. While classical ML algorithms are still used for EL or PL image analysis [74–77], the community is increasingly shifting towards using advanced deep learning algorithms for more accurate results.

Defect classification, detection, and segmentation are among the primary uses of deep learning. The distinction between these tasks is illustrated in Fig. 1. Defect classification determines whether a defect is present in a solar cell, while defect detection provides the location of the defect(s) with bounding boxes. Lastly, defect segmentation is a more precise method of locating defects, as it identifies the exact location and boundaries of each defect within the cell.

Identifying defects during or after cell processing, but before module manufacturing, is crucial to prevent performance and reliability issues in installed PV systems. Usual defects, such as cracks and finger failures, arise during cell manufacturing. Other defects, such as potential or light-induced degradation, delamination, glass breakage, or soiling, are caused by environmental stress and can be identified through module EL or IRT imaging.

This section explores the various applications of deep learning in defect classification, detection, and segmentation (mainly on luminescence images), as well as other novel applications.

3.1. Defect classification

Deep learning algorithms can be used on classification tasks such as identifying whether defects are present or not. For the training of such algorithms, a labelled dataset is required, where each instance is assigned to a specific class. Generally, larger labelled datasets (>1,000) are preferred as they lead to better algorithm performance and reduced overfitting error. However, creating such large, labelled datasets can be a time-consuming and expensive process.

A dataset containing 2,654 mc-Si and mono-Si EL images, extracted from 44 modules, was published in 2018 by Buerhop-Lutz et al. [78]. The dataset, referred to as the ELPV dataset, is available on Github [78] and each image in the dataset is 300 × 300 pixels and is annotated as Functional or Defective.

The ELPV dataset was first used by Deitsch et al. [79] for defect classification. They employed two approaches for defect classification using this dataset: a deep learning-based convolutional NN (CNN) model (VGG) [80] and a ML-based model (SVM) [81]. Their study demonstrated the superiority of automated feature extraction through a CNN compared to handcrafted features. The CNN model achieved an F_1-score of 88.4 %, while the non-deep learning approach involving feature engineering and an SVM classifier achieved an F_1-score of 82.4 %. Several groups have since utilised the ELPV dataset for the comparison of deep ML models. Akram et al. [82] introduced data augmentation techniques, such as random horizontal and vertical flipping, and contrast operations to increase the size of the ELPV dataset. Their architecture was also based on the VGG CNN with a reduced number of layers and achieved an F_1-score of 92.2 %, with data augmentation shown to increase the accuracy by 6.5 %. Another study by Acharya et al. [83] relabelled the dataset into four classes ('no defect', 'micro defect', 'large defect', and 'low-resolution defect') and trained a custom CNN architecture achieving a 73.8 % F_1-score, outperforming the results of a VGG16 classifier (65.0 % F_1-score) on the same dataset. Their custom CNN was based on a Siamese CNN [84], where the left and right side of the EL image were passed to separate twin pipelines of the CNN.

Ge et al. [85] proposed a hybrid fuzzy CNN approach that improved the F_1-score from 80.0 % (VGG) to 82.0 %. However, they suggested that the performance could be further improved with larger datasets. Demirci et al. [86] combined the CNN feature extractor and SVM classifier for binary classification on the ELPV dataset. By using multiple CNN architectures and reducing the number of features, they achieved an F_1-score of 95.5 %. Huang et al. [87] compared different CNN architectures for binary classification and introduced a swarm-based pruning method that reduced the memory usage by 100 times while matching the performance of the best CNN architecture. This enables the deployment of the models on mobile and embedded devices.

In a span of three years, deep learning-based algorithms for defect classification improved with a learning curve from 88.4 % to 95.5 % F_1-score. These values are higher compared to the performance of the SVM classifier using custom features for the same task with an 82.4 % F_1-score. A summary of the results on the ELPV dataset is shown in Table 7.

In a study by the ELPV publishing group, Bartler et al. [88] expanded the ELPV dataset from 44 to 1,366 modules. However, this resulted in an imbalanced dataset of 98,280 cells with only 3.4 % defective cells. To address this issue, they applied data augmentation techniques, such as random horizontal and vertical flipping, and resampling. The authors then trained a VGG-based CNN network to classify the cells, achieving an impressive F_1-score of 93.0 %, which represents a significant improvement over the work of Deitsch et al. [79]. These results emphasise the importance of large and balanced datasets to effectively train deep learning algorithms. However, it is worth noting that the authors did not publish the extended dataset, which could limit further research in the field.

The results on larger datasets encouraged the use of more complex deep learning algorithms. For example, Ying et al. [89] used 13,835 mc-Si cell EL images to train a custom CNN architecture as a feature extractor and an RF to classify into six classes: 'no defects', 'open welding', 'broken grid', 'solid black', 'shadow', and 'hidden crack'. The authors trained a multi-channel CNN where the EL image was resized for each of the channels capturing simple to more complex features in each channel achieving an F_1-score of 93.0 %. Karimi et al. [90] applied a CNN-based technique for defect classification of cells in PV modules that had undergone different lengths of damp heat exposure, introducing

Table 7
Defect classification performance on the ELPV dataset.

Reference	Year	ML model	Classes	F_1-score
Deitsch et al. [79]	2019	Custom features + SVM CNN(VGG) + SVM	functional; defective	82.4 % 88.4 %
Akram et al. [82]	2019	CNN(VGG)	functional; defective	92.2 %
Acharya et al. [83]	2020	CNN(Custom) CNN(VGG)	no defect; micro defect; large defect; low resolution defect	73.8 % 65.0 %
Ge et al. [85]	2021	CNN(VGG) CNN(Hybrid fuzzy)	defective; maybe defective; maybe functional; functional	80.0 % 82.0 %
Demirci et al. [86]	2021	CNN(VGG) + SVM CNN(Ensemble) + SVM	functional; defective	91.9 % 95.5 %
Huang et al. [87]	2022	CNN(VGG) CNN(AlexNet) CNN(ResNet) PSO-Pruned CNN	functional; defective	91.1 % 91.8 % 94.8 % 94.2 %

Table 8
Defect classification performance on other datasets.

Reference	Year	ML model	Dataset	Classes	Performance
Bartler et al. [88]	2018	CNN(VGG)	ELPV (extended)	functional; defective	F_1-score: 92.2 %
Ying et al. [89]	2018	CNN(custom) + RF	13,835 mc-Si cells from module EL	no defect; open welding; broken grid; solid black; shadow; hidden crack	F_1-score: 93.0 %
Karimi et al. [90]	2019	CNN(custom) RF SVM	3,550 mc-Si and mono-Si cells EL	good; cracked; corroded	F_1-score: 98 % F_1-score: 96 % F_1-score: 82 %
Luo et al. [91]	2019	CNN(SqueezeNet) CNN(AlexNet) CNN(ResNet)	507 defective monoSi EL into 40,000 GAN generated images	grid fingers; material defect; microcrack; deep cracks	Accuracy: 60.6 % Accuracy: 74.7 % Accuracy: 98.4 %
Tang *et al.* [92]	2020	CNN(VGG) CNN(Custom)	ELPV + Jinko 1,800 EL images with GAN augmentation: 7,400	defect-free; micro-crack; finger-interruption; break	Accuracy: 82 % Accuracy: 82.5 %
Verma et al. [93]	2021	CNN(AlexNet)	392 module EL images with geometric transformations for data augmentation	functional; defective	Accuracy: 85.2 %
Et-taleby et al. [94]	2022	CNN (VGG-16) with SVM	Dataset 1–2,624 EL images Dataset 2–1,028 EL images	D1 – good; defective D2 – good; corroded/cracked	Accuracy: D1 – 99.49 % D2 – 99.46 %
Fioresi et al. [73]	2022	CNN(ResNet)	17,064 El images	cracks, front grid defects, interconnect defects, contact defects	F_1-score: 95.0 %

cracks and corroded defects. The authors showed that their custom CNN achieved 96 % accuracy compared to a direct classification with RF (96 %), and SVM (82 %). The F_1-scores were not reported, although the F_1-score is more indicative of model performance for imbalanced datasets compared to accuracy.

To increase dataset sizes and increase the defective to non-defective ratio, geometric transformations for data augmentation are commonly used [82]. However, generative adversarial networks (GAN) can also be trained to generate diverse samples from a few examples of a defect class. With GANs, Luo et al. [91] generated a dataset of 40,000 images for each of the four classes ('Grid fingers, 'Micro-crack', 'Material defect', and 'Deep crack') from only 507 defective mono-Si EL images. Multiple CNN architectures were then trained for the classification task and achieved an accuracy of 98.42 % (ResNet), 74.73 % (AlexNet), and 60.6 % (SqueezeNet). Although the F_1-scores were not reported, they are expected to be close to the accuracies due to the balanced nature of a GAN-generated dataset. However, the authors mentioned that due to disproportionately higher numbers of the original datasets toward the "Grid fingers" class, the trained CNN achieved lower accuracy for the other classes. This can be explained by the limited diversity of defect examples generated by the GAN for classes with fewer examples. Tang et al. [92] also used GAN for data augmentation and trained a VGG CNN to classify four classes with 82 % accuracy.

Verma et al. [93] used geometric transformations for data augmentation and analysed the impact of learning rate and number of epochs on the training of an AlexNet CNN. They achieved accuracies of up to 85.2 %. Slower learning rate and longer training time are recommended by the authors. However, the F_1-scores are not reported. Et-taleby et al. [94] applied a CNN (VGG-16) along with an SVM to two datasets containing 2,624 EL images of mono-Si and mc-Si, and 1,028 EL images of commercial modules from various brands. Their results show that the accuracy scores of their developed model on the first and second datasets are 99.49 % and 99.46 %, respectively. Fioresi et al. [73] introduced a publicly available annotated dataset of 17,064 EL images, training a CNN for defect classification with a high F_1-score of 95 %. Studies on defect classification on other datasets are summarised in Table 8.

The ELPV dataset provides a clear way to benchmark new approaches or deep learning algorithms as they improve year on year, reaching a 95.5 % F_1-score. For multi-classification task on other datasets, CNN based approaches improved both F_1-score and accuracies since 2018 with the implementation of data augmentation strategies using GANs.

3.2. Defect detection

While classification models identify whether defects are present in a solar cell luminescence image, defect detection models specifically localise these defects within the images. Hence, defect detection identifies the parts of image that contain defects. The goal of this detection algorithm is to output a bounding box, or region of interest, around the defects in the image.

Defect detection algorithms are usually built on the region-based CNN (R–CNN), a slow but more accurate model, or the you-only-look-once (YOLO) framework, traditionally faster but more prone to errors. As the R–CNN has been updated multiple times [95,96], the Faster R–CNN version has been used in most of the literature reviewed in this section. Liu et al. [97] compared different potential backbones of the Faster R–CNN for the application of surface defect detection in mc-Si EL images. They concluded that deep CNN structures, such as ResNet-101, perform the best with a mean average precision (mAP) of 88.34 %. They also improved the Faster R–CNN by adding a feature pyramid network (FPN) to increase the robustness of the feature extraction, achieving a mAP of 89.97 % on their validation set. They also applied a guided anchoring region proposal network (GAR) which is used to adapt the shape and position of the bounding box to the feature of the defect, further increasing the mAP to 94.62 %. Another popular deep learning-based detection algorithm, YOLO, is used by Zhang et al. [98] for detecting broken, unsoldered or cracked cells in mc-Si cell EL images from modules. They conclude that YOLOv3, with a mAP of 82.5 %, slides in between the Faster R–CNN (80.1 %) and the Faster R–CNN + FPN (82.9 %). They also introduced a fused multi-channel CNN between the R–CNN and its FPN counterpart to achieve a mAP of 85.7 %.

To address issues of low detection precision of certain features, data-mining strategies, such as gathering online EL images, were employed. A dataset expansion using Gaussian noise filtering and image binarisation was implemented by Xu et al. [99], generating up to 11,450 defect images from 870 mono-Si cell EL images (unpublished dataset). The authors propose the use of a single-shot detection (SSD) algorithm built

Table 9
Defect detection performance on luminescence images.

Reference	Year	Model	Dataset	Detection of	mAP	FPS/(f.s^{-1})
Liu et al. [97]	2019	Faster R–CNN (ResNet) Faster R–CNN + FPN Faster R–CNN + FPN + GAR	5,000 mc-Si cell EL	surface defects	88.34 % 89.97 % 94.62 %	0.83 2.94 5.26
Zhang et al. [98]	2020	YOLOv3 Faster R–CNN Faster R–CNN + FPN Multi-channel R–CNN	mc-Si cell EL from 1,462 modules and online data mining	broken, unsoldered, cracked	82.5 % 80.1 % 82.9 % 85.7 %	N/A 2.36 2.59 1.57
Xu et al. [99]	2021	YOLOv3 Faster R–CNN Improved SSD	870 mono-Si cell EL images	scratch, broken gate, dark, pollution	62.2 % 68.2 % 73.4 %	N/A N/A 0.56
Su et al. [100]	2021	YOLOv3 Faster R–CNN Faster R–CNN + BAF	3,629 mono-Si cell EL images in 1/8th patches	cracks, finger failure, black core	78.79 % 83.11 % 88.07 %	11.36 6.41 7.75
Su et al. [101]	2021	YOLOv3 Faster R–CNN Faster R–CNN + CAN	3,629 mono-Si cell EL images in 1/8th patches	cracks, finger failure, black core	78.79 % 83.11 % 87.38 %	N/A 9.16 9.10
Meng et al. [72]	2022	YOLOv4	2,144 half-cell module EL images	black area, cracks, break, finger failure, low cell, scratch, black cell	94.55 %	35
Li et al. [102]	2023	GBH-YOLOv5	307 half-cell module EL images	no electricity, hot spot, scratch, black border, broken	97.8 %	1.70
Huang et al. [103]	2023	SSD Faster R–CNN YOLOv5 YOLOv5-BDL	4,964 mono-Si cell EL images	crack, finger, black core, material, short circuit, thick line, horizontal dislocation	92.2 % 73.18 % 93 % 95.5 %	17.01 7.27 80.65 53.92

on a VGG-16 and FPN, achieving a mAP of 73.4 %, outperforming the classic Faster R–CNN (68.2 %) and YOLOv3 (62.2 %). Contrary to Zhang et al.'s [98] study, the Faster R–CNN performs better than YOLOv3 on that particular dataset.

Two studies from Su et al. [100,101] drove the performance even higher with a couple of innovations in the deep learning model's structure. Both studies used the same dataset of 3,629 mono-Si cell EL images (unpublished) containing a high density of images of defects (~59 %) such as cracks, finger failure, and black core. Each EL image (1024 × 1024) is split into eight separate patches (128 × 128) to enhance the size of the dataset, keeping the high-resolution of defects and lowering the training time. In their first study [100], they upgraded the FPN approach with a bi-directional attention feature pyramid network (BAF) further moving the network's attention away from non-important background features to relevant defect features. With the Faster R–CNN + BAF, they achieved a mAP of 88.07 % compared to 78.79 % for YOLOv3 on the same dataset. The improvement over Faster R–CNN for the BAF (+4.96 %) is significantly better than for the FPN (+2.5 %) from Zhang et al. [98]. In their second study [101], they improved their model by integrating a complementary attention network (CAN) for adaptive background feature suppression and defect feature highlighting, reaching an mAP of 87.38 %, an improvement of 4.28 % over the Faster R–CNN. The Faster R–CNN from Su et al. [100, 101] performs significantly better (mAP: 83.11 %) than YOLOv3 (mAP: 78.79 %) agreeing with Xu et al.'s [99] finding. While the later applications of Faster R–CNN achieve incredible prediction accuracy, the above models all have prediction speeds below 10 frames per second [fps (f.s^{-1})], which may cause challenges if faster applications are needed. Nonetheless, advancements in computer technology could offer substantial enhancements to the computational speeds of AI systems.

Meng et al. [72] compared the use of multiple backbone CNN on the YOLOv4 framework (maximum of mAP: 94.55 %) to detect defects on half-cell EL images, highlighting the use of deep learning detection algorithms on other cell morphology and achieving a very fast prediction time of 35 fps. Similarly, Li et al. [102] utilised the YOLOv5 framework with many additional modules such as a Ghost convolution and BottleneckCSP followed by an FPN and path aggregation network. The dataset used contained 307 EL images of half-cell modules labelled with five different defects, and the trained model achieved an average mAP of 97.8 % but sacrificed the inference time to just 1.7 fps. Their annotated dataset has been shared and other studies can use the dataset to benchmark their models. Huang et al. [103] also utilised an improved YOLOv5 with a bi-directional FPN, achieving both improved mAP (95.5 %) compared to a basic Faster R–CNN (73.18 %), while still maintaining a fast prediction speed of 53.92 fps.

The results of the deep learning algorithms for defect detection are shown in Table 9. When comparing prediction scores (mAP), most of the R–CNN architectures outperform the YOLO-based models. However, the speed of the models is also an important factor. As YOLO-based models are often much faster compared to R-CNNs, a user may choose to optimise for speed or prediction accuracy depending on the context of the defect detection application. Incidentally, the development of high-performing defect detection algorithms can lead to high-performing algorithms for defect classification tasks, as shown by Su et al. [101] (F_1-score = 98.47 % with their Faster R–CNN + CAN). Like the previous defect classification sub-section, a lack of published datasets constrains the authors to compare all the different algorithms on their dataset, making it more difficult to compare to the rest of the literature.

3.3. Defect segmentation

Another deep learning approach to localise defects in a luminescence image is through semantic segmentation, i.e., a pixel-level classification task where the algorithm is tasked to find all the pixels in the image corresponding to a defective area.

Mayr et al. [104] trained a ResNet CNN for a classification task on the ELPV dataset as a surrogate task for segmentation. The authors achieved an F_1-score of 83 % for the classification, lower than the CNN models of Deitsch et al. [79] and Akram et al. [82] from the same year. However, the authors' approach to the ResNet CNN differs from more classic CNN models as they modified the convolutional layers and removed the pooling layers to preserve the spatial information through the forward

path of the CNN. The fully connected layer, which performs the classification into the different bins, is replaced with a convolutional layer, effectively outputting an activation map per classification bin. After training for the classification task, a thresholding method on the CNN's activation map performs the segmentation.

Rahman et al. [105] trained a U-Net with a global attention block (GAB) for each U-Net level connection to weigh the important image features. The segmentation training of the U-Net was performed on 400 cell EL images with defects, achieving an intersection-over-union (IoU) of 0.630. The addition of the GAB increased the performance to 0.648. Another type of attention technique was introduced by Jiang et al. [106], with the addition of the popular squeeze and excitation blocks (SE) which improved the accuracy and precision of the segmentation. The model was trained on a subset of ELPV (25 mono-Si and 25 mc-Si EL images with defects), achieving an IoU of 0.661 and a Dice coefficient of 0.480 without the attention SE blocks, and an IoU of 0.693 and a Dice of 0.541 with the attention SE blocks.

Qian et al. [107] implemented a segmentation approach combining CNN-based features and a stacked denoising AE. The model was trained on the ELPV dataset achieving a pixel-level F_1-score of 46.88 %. Similarly, Otamendi et al. [108] used a CNN for classification and autoencoder (AE) for the segmentation task of defects, achieving a structure similarity index (SSIM) of 0.992. The dataset used was a combination of the ELPV dataset and an additional 5,592 cell EL images from the authors' industrial partners, including half-cut cell images. The additional images included different cell morphology such as heterojunction cells and half-cut cells, showing the potential application of deep learning techniques across the different PV technologies.

More use cases of implementation of segmentation were performed by Pratt et al. [109], using a wide variety of cell EL images extracted from modules from manufacturing and post thermal and mechanical stress testing. The authors trained a U-Net to generate pixel-wise classification for multiple classes of defects, such as cracks and inactive areas. Despite the low number of images in the dataset (150 cell images including 30 in the validation set), the authors show qualitative agreement between the ground truth and predicted segmentation across the wide variety of cells in the dataset.

Tian et al. [110] compared three algorithms for the segmentation task on the ELPV dataset: CNN, U-Net, and U-Net++, a more complex U-Net with densely connected nested decoder sub-networks. The authors showed that the U-Net++ performed best with an IoU of 0.955 compared to the U-Net (0.929) and the CNN (0.892). Jiang and Zhao [111] also applied a dual U-Net, one to output cell classification, with near-perfect accuracy, and another for defect segmentation, on the ELPV dataset, reaching an IoU of 0.693 with the use of a CAN. This solidifies the trend toward U-Net as the framework of choice for segmentation tasks.

Wang et al. [112] proposed an efficient and refined deep convolutional features network (ERDCF) which uses side outputs from an encoder to refine the decoder of a CNN. The authors claim this helps with background noise and reduces the computational complexity by limiting the size of the feature map. This is confirmed on their private dataset where they compared their network, achieving an IoU of 0.936, to a classic U-Net structure (IoU: 0.865). This shows there is still design space to develop new architecture to improve on the segmentation problem of defects in luminescence images.

A different approach for defect segmentation is taken by Ni et al. [113], looking to reconstruct defective EL images into defect-free images. The image difference then highlights the defective area and completes the segmentation task. The authors trained a deep belief network (DBN) for the reconstruction process on mc-Si cell EL images (dataset unpublished). An improvement on the reconstruction algorithm by implementing GANs is introduced by Buratti et al., [114] where the authors used the difference image between an original and a GAN-reconstructed image to highlight the defect, effectively segmenting the defective area of an EL image. The difference between reconstructed and original images was also used to assess the efficiency impact of the presence of the defect. Su et al. [115] used a CycleGAN network and introduced a strong identity loss to translate images between defective and defect-free domains, achieving an IoU of 0.887. Balzategui et al. [116] introduced an automatic labelling deep learning approach based on GANs. Their approach enables the detection of anomalous patterns on solar cells as unsupervised learning (no labelling needed). This is based on the training of GAN on non-anomalous data and uses the difference in image between the original and reconstruction images as a segmentation and automatic annotation and labelling of the image. The performance comparison between the different reviewed studies for defect segmentation can be found in Table 10.

Based on the reviewed studies, the U-Net architecture seems to be the preferred algorithm for semantic segmentation, while the image

Table 10
Defect segmentation performance on luminescence images.

Reference	Year	ML model	Dataset	Metric	Performance
Ni et al. [113]	2018	DBN + Image difference	540 mc-Si EL images	Qualitative	–
Mayr et al. [104]	2019	CNN activation map with surrogate classification task	ELPV	Qualitative	–
Rahman et al. [105]	2019	U-Net U-Net + GAB	400 cell EL images with defects	IoU	0.630 0.648
Jiang et al. [106]	2020	U-Net U-Net + SE	Subset of ELPV	IoU | Dice	0.661 | 0.480 0.693 | 0.541
Qian et al. [107]	2020	CNN + AE	ELPV	Pixel F_1-score	46.88 %
Otamendi et al. [108]	2021	AE	ELPV and 5,592 EL images	SSIM	0.992
Pratt et al. [109]	2021	U-Net	150 cells EL images from stressed modules	Qualitative	–
Tian et al. [110]	2021	CNN U-Net U-Net++	ELPV	IoU	0.892 0.929 0.955
Buratti et al. [114]	2021	GAN + Image difference	30,000 cell mono-Si EL images	Qualitative	–
Su et al. [115]	2021	GAN + Image difference	540 cell mc-Si EL images	IoU	0.887
Balzategui et al. [116]	2021	GAN + Image difference	1873 cell mono-Si EL images	Qualitative	–
Jiang and Jao [111]	2022	U-Net U-Net + CAN	ELPV	IoU	0.661 0.693
Wang et al. [112]	2022	U-Net ERDCF	300 cell EL with cracks	IoU	0.865 0.936

difference approach, mainly with the use of GAN for the reconstruction step, is also promising. A future step would be to implement instance segmentation, where the algorithms need to separate the different instances of defects instead of just the defect class. However, there is a lack of agreement on the performance metric for segmentation, compared to detection (mAP) and classification (F_1-score). IoU is the most reported metric and should be systematically reported in future segmentation studies. Furthermore, there is a need for published datasets for segmentation, just as in detection and classification, to serve as a baseline for future development and comparison with the literature. This is particularly important for segmentation, as it is more time-consuming and reliant on experts.

3.4. Other deep learning applications

In a study by Buratti et al. [117], a two-step deep learning framework was introduced for efficiency binning and efficiency prediction of solar cells. The approach was applied to full and half-cut cells, demonstrating the potential for luminescence imaging to replace I–V measurements in PV manufacturing lines. Although the data is not published, the trained algorithm models were shared for future transfer learning and fine-tuning applications. The same authors presented another study [114] where GANs were trained to reconstruct the defective part of EL images, and the developed efficiency prediction capability was used to assess the efficiency pre and post GAN-reconstruction, effectively implementing an automated loss-analysis with the use of deep learning on luminescence images. In a different study, Kunze et al. [118] developed a deep learning pipeline combined with an empirical model to estimate the I–V parameters of a solar cell from EL, PL images, and spectral reflection measurements, including the generation of the I–V curves and cell binning.

Deep learning methods have also been applied to post-process EL images. Dwivedi et al. [119] used enhanced super-resolution GAN (ESRGAN) to improve EL image resolution, allowing for a more detailed analysis of these images. Similarly, Liu et al. [120] employed U-Net to effectively denoise EL images and enhance the overall image quality such that cheaper hardware or faster exposure times could be utilised.

Deep learning applications from luminescence images also branch out along the value chain of PV manufacturing. On bare Si wafers (pre-processing solar cell), Demant et al. [121] introduced a DenseNet-based CNN to predict the expected I–V characteristics of the full solar cell from the as-cut mc-Si wafers' PL images. The authors extended their analysis of the PV production line using deep learning with dataset quality visualisation in a following study [122]. A third study [123] by the same authors introduced GANs for denoising the wafer images, further improving the prediction capability of the CNN. Kurumundayil et al. [124] also used GAN for denoising diamond wire sawed wafers. Fu et al. [125] used CNNs to predict the incident photon-to-electron conversion efficiency of mono-like cast Si from PL images of bare silicon wafers. Han et al. [126] implemented R–CNN and U-Net for crack segmentation of bare wafer luminescence images. For optimising designs of solar cells, Bhattacharya et al. [127] developed SolarNet, a CNN-based power optimisation tool, combined with solar cell physics equations to design metallisation patterns in the two-dimensional domain.

On the other side of the PV manufacturing value chain, Karimi et al. [128] used deep learning on module EL images and time-series I–V measurements to predict PV module power and series resistance when the modules are undergoing thermal cycling stresses. The authors demonstrated the capabilities of deep learning in predicting degradation mechanisms in three different types of Si cells. Similarly, Hoffman et al. [129] developed a deep learning-based method for determining the power at maximum power point of Si PV modules from on-site or indoor EL images. The dataset, containing 719 EL images at various stages of degradation from 137 mono-SI and mc-Si modules, was published by the authors, under the name of ELPVPower [129] and can be used for reference in a future study involving module EL image analysis or degradation prediction.

Apart from luminescence imaging, deep learning techniques have also been used in different imaging techniques. Kunwar et al. [130] implemented CNNs on synchrotron radiation images in PV ribbon fabrication. Their goal was to automatically detect interfacial microstructures from thousands of images. Li et al. [131] implemented defect classification on images in the visible light range taken by a drone flying over a PV power plant looking for defects such as dust shading, encapsulant delamination, glass breakage or yellowing. Similarly, Chen et al. [132] looked at the visible light range of cells for defect classification using deep learning. Most of the non-luminescence work resides on the analysis of IRT with deep learning classification methods done by Pierdicca et al. [133], Akram et al. [134], and Bommes et al. [135] on module images, each with an increased complexity in the classification algorithms. A dataset of 20, 000 labelled IRT images of modules with 11 classes was recently published, along with a defect classification implementation by Fonseca et al. [136]. Zefri et al. [137] applied a deep learning framework for aerial inspection from drones using IRT images. Su et al. [138] applied their attention mechanism innovation for the detection of hot spots on IRT images. Finally, at the other end of the spectrum, Gilleland et al. [139] implemented an SSD for defect detection on UVF images.

4. Conclusions and recommendations

This review provides a comprehensive analysis of recent studies that have implemented ML techniques for characterising Si-PV devices. Specifically, the application of classical ML techniques for optimising processes, investigating devices, and assessing manufacturing lines, as well as deep learning techniques for classifying, detecting, and segmenting defects along with other niche applications are examined.

4.1. Data management

Data plays a critical role in any ML application, including PV characterisation. The size of the dataset is key in developing accurate ML models as larger datasets generally provide more reliable results. Publicly available datasets are important as they provide a benchmark for comparison and encourage further development. To ensure the ease of use of these datasets, it is recommended to use data and metadata storage principles. One such principle is FAIR, which focuses on sharing data such that it is Findable, Accessible, Interoperable, and Reusable [140]. These guidelines emphasise the need for comprehensive documentation, machine-readable data/metadata, and accessible computational resources.

Additionally, protocols should be clearly defined, specifying the format and the required data. For example, in solar cell production lines, standard guidelines should be implemented for naming equipment parameters, specifying measurement units, and detailing data recording practices. This would facilitate consistent comparisons across different production lines. Adopting uniform terminology across the industry, potentially starting with equipment manufacturers, could improve clarity and operational coherence.

Data processing also plays a crucial role. Labelling datasets can be labour-intensive and prone to errors. This process can be streamlined using unsupervised algorithms to automate labelling. Additionally, techniques like geometric transformations or GAN-based data augmentation can be used to effectively increase dataset size.

Currently available datasets on Si-based PV are shown in Table 11. While it is ideal to publish datasets, researchers should aim to at least make their trained models available for transfer learning and fine-tuning applications if sharing the datasets is not possible.

Table 11
Publicly available datasets.

Reference	Year	Data	Link
Buerhop-Lutz et al. [78]	2018	2,624 EL images of cells	https://github.com/zae-bayern/elpv-dataset
French et al. [141]	2019	EL images of three 60-cells modules at different dampheat degradation	https://osf.io/v6pwe
Hossain et al. [69]	2019	400 I–V of cells	https://figshare.com/articles/dataset/Dataset_-_A_Comprehensive_Methodology_to_Evaluate_Losses_and_Process_Variations_in_Silicon_Solar_Cell_Manufacturing/7636922
Hoffmann et al. [129]	2020	719 EL images of modules at various stages of degradation	https://github.com/ma0ho/elpvpower
Millendorf et al. [142]	2020	20,000 IR images of modules	https://github.com/RaptorMaps/InfraredSolarModules
Fioresie et al. [73]	2022	17,064 EL images of cells	https://github.com/ucf-photovoltaics/UCF-EL-Defect
Li et al. [102]	2023	307 half-cell module EL images	https://github.com/CCNUZFW/PV-Multi-Defect/tree/main

4.2. Model evaluation and validation

Different algorithms were applied to tasks like regression, optimisation, classification, and outlier detection. In regression tasks, NNs are the most common models. In optimisation tasks, parameter sweeps are usually done, however, more advanced algorithms such as GA and PSO are recommended. In classification tasks like defect classification, CNN-based architectures are extensively implemented. For defect detection, Faster R–CNN stands out as the algorithm of choice in the reviewed studies. Numerous algorithm improvements of the Faster R–CNN have been explored in the last years, however, advancements in YOLO (v5 – v7) for improved accuracy combined with its inherent faster prediction speed suggest that future PV defect detection models may begin to rely heavily on the YOLO algorithm. Better defect detection algorithms directly impact the quality of defect classification algorithms. For defect segmentation, either a GAN-based image difference approach or a pixel-level classification U-Net approach is recommended.

There is a need to use uniform metrics across studies to help the PV community easily compare ML models. For optimisation, most studies did not report a quantified performance of their approach but the RMSE or overall performance increase from the non-optimised to optimised state is recommended. For regression, varying metrics were reported but RMSE and R^2 are recommended. For classification, a confusion matrix and the F_1-score for each class are preferred to easily identify the overall performance of the model and the problematic classes. The F_1-score is preferred over the accuracy as it is not influenced by class imbalance. For detection and segmentation, the recommended metrics are mAP and IoU, respectively. Examples of successful and unsuccessful detection and segmentation should also be shown.

Model validation is also key in ensuring the robustness of ML models. This can include testing the trained models on unseen data from simulations and/or experimental measurements. Lastly, discussing areas for improvement is helpful in guiding future studies to enhance their approach.

4.3. Classical ML vs deep learning

In Section 3, it was shown that deep learning algorithms perform better than traditional ML algorithms (RF, NN, SVM). The development of deep learning algorithms can be applied to data-driven modelling and characterisation applications discussed in Section 2. A simple way to access the suite of CNN-based algorithms is to transform a linear regression or classification problem from Section 2 into an image-based problem. This can be done when more than one key variable is present as the predictive target can be plotted as a bi-dimensional map, either with PCA or as multiple bi-dimensional slices.

Deep learning is widely used in various image analysis tasks. However, classical ML also maintains an important position in the pipeline, typically being utilised towards the end. Both deep learning and classical ML techniques offer distinct roles and advantages depending on the application. Classical ML is often employed for straightforward predictive modelling and classification, particularly with smaller datasets. In contrast, deep learning excels in complex pattern recognition tasks, such as feature extraction from large image datasets. These techniques can complement each other in certain applications, and studies have shown benefits from integrating both within the same application [94,117].

4.4. Outlook

Despite significant advancements in ML applications in PV, future research should focus on integrating these models into PV research and industrial settings. Within the PV research community, a clearer grasp has been obtained on how AI can be specifically applied, which allows a greater focus on application optimisation rather than on basic algorithmic advancements. However, future challenges for the deployment of these models include model interpretability (understanding their "black box" nature), prediction speed versus accuracy, overall computational demands of the application, and integrating niche applications into a unified framework. Other fields have started tackling similar challenges through techniques like explainable AI [143] and ablation studies [144]. Addressing these issues builds upon the robust knowledge base of ML in PV and will assist the deployment of ML models into different areas of the PV industry.

Interdisciplinary collaboration is also critical for the application of ML to PV characterisation. A combination of technical skills from PV, computer, and data scientists and engineers is needed to produce robust ML applications for the PV industry. This is evident from the diverse technical backgrounds of the authors cited within this review. PV conferences should foster more collaborative environments by improving the integration of ML topics. The same effect may also be achieved by inviting PV studies to be presented at computer science-based conferences. Individual institutions and industries should also foster environments for different scientists and industry representatives to meet, such as focus hubs and dedicated collaborative umbrella institutes (e.g. UNSW just established its AI Institute that aims to deploy AI responsibly across different disciplines). Exploring methods to encourage industry involvement necessitates addressing potential concerns, such as data confidentiality, which could be possibly mitigated through techniques like data normalisation. By creating spaces for collaboration and knowledge sharing, the application of ML to all aspects of PV research and development can continue to flourish.

The application of ML to PV characterisation has the potential to unlock new levels of performance and reliability of cells and modules and improve manufacturing throughput, helping to achieve the targets set by the ITRPV and the IPCC. This would also lead to increased clean energy production from PV due to the higher power and reliable PV devices being manufactured with the aid of ML-powered characterisation. For example, ML algorithms can identify and mitigate material defects early in the manufacturing process, leading to higher-performing cells/modules with extended lifetimes and improved energy production

potential. While the range of ML applications covered in this study is impressive, it is clear these are only scratching the surface of what is possible with ML in the PV industry. As the field continues to evolve and mature, even more innovative and impactful uses of ML and other data-driven techniques in PV research and development are expected.

Authors contribution

Y.B., G.M.N.J., and Z.A.V. contributed equally as the first authors in the writing of the initial version of the manuscript. G.M.N.J. and Z.A.V. handled the manuscript submission and subsequent revisions. P.D and Z. H contributed to the discussion and direction of the paper. Z.H also supervised the work.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Zubair Abdullah-Vetter reports financial support was provided by Australian Renewable Energy Agency.

Data availability

Publicly available data has been linked in the manuscript.

Acknowledgment

This work was supported by the Australian Government through the Australian Renewable Energy Agency [ARENA; Projects 2020/RND016 and 2022/TRAC002]. The views expressed herein are not necessarily the views of the Australian Government, and the Australian Government does not accept responsibility for any information or advice contained herein.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rser.2024.114617.

References

[1] Masson-Delmotte V, Zhai P, Pörtner H-O, Roberts D, Skea J, Shukla PR, et al. Global warming of 1.5°C - an IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways. In: The context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. In Press: IPCC; 2018.
[2] Perez R, Perez M. A fundamental look at energy reserves for the planet. The IEA SHC Solar Update 2015;50:2–3.
[3] IEA. Net zero by 2050. Paris: IEA; 2021.
[4] Fisher M, Woodhouse M, Herritsch S, Trube J. International technology roadmap for photovoltaics. twelfth ed. 2021. VDMA e. V.
[5] Fuyuki T, Kondo H, Yamazaki T, Takahashi Y, Uraoka Y. Photographic surveying of minority carrier diffusion length in polycrystalline silicon solar cells by electroluminescence. Appl Phys Lett 2005;86:262108. https://doi.org/10.1063/1.1978979.
[6] Trupke T, Bardos RA, Schubert MC, Warta W. Photoluminescence imaging of silicon wafers. Appl Phys Lett 2006;89:044107. https://doi.org/10.1063/1.2234747.
[7] Ebner R, Kubicek B, Újvári G. Non-destructive techniques for quality control of PV modules: infrared thermography, electro- and photoluminescence imaging. In: Iecon 2013 - 39th annual conference of the IEEE industrial electronics society; 2013. p. 8104–9. https://doi.org/10.1109/IECON.2013.6700488.
[8] Morlier A, Siebert M, Kunze I, Mathiak G, Köntges M. Detecting photovoltaic module failures in the field during daytime with ultraviolet fluorescence module inspection. IEEE J Photovoltaics 2017;7:1710–6. https://doi.org/10.1109/JPHOTOV.2017.2756452.
[9] Fisher M, Woodhouse M, Herritsch S, Trube J. International technology roadmap for photovoltaics. twelfth ed. 2021. VDMA e. V.
[10] El Chaar L, lamont LA, El Zein N. Review of photovoltaic technologies. Renew Sustain Energy Rev 2011;15:2165–75. https://doi.org/10.1016/j.rser.2011.01.004.
[11] Vicari Stefani B, Kim M, Zhang Y, Hallam B, Green MA, Bonilla RS, et al. Historical market projections and the future of silicon solar cells. Joule 2023;7:2684–99. https://doi.org/10.1016/j.joule.2023.11.006.
[12] Wolff J, Pauling J, Keck A, Baumbach J. The economic impact of artificial intelligence in health care: systematic review. J Med Internet Res 2020;22:e16866. https://doi.org/10.2196/16866.
[13] Dimiduk DM, Holm EA, Niezgoda SR. Perspectives on the impact of machine learning, deep learning, and artificial intelligence on materials, processes, and structures engineering. Integr Mater Manuf Innov 2018;7:157–72. https://doi.org/10.1007/s40192-018-0117-8.
[14] Tuomi I, Cabrera M, Vuorikari R, Punie Y, European Commission, Joint Research Centre. The impact of artificial intelligence on learning, teaching, and education: policies for the future. 2018.
[15] Mellit A, Kalogirou SA. Artificial intelligence techniques for photovoltaic applications: a review. Prog Energy Combust Sci 2008;34:574–632. https://doi.org/10.1016/j.pecs.2008.01.001.
[16] Yang D, Kleissl J, Gueymard CA, Pedro HTC, Coimbra CFM. History and trends in solar irradiance and PV power forecasting: a preliminary assessment and review using text mining. Sol Energy 2018;168:60–101. https://doi.org/10.1016/j.solener.2017.11.023.
[17] 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. https://doi.org/10.1016/j.rser.2017.04.046.
[18] AbdulMawjood K, Refaat SS, Morsi WG. Detection and prediction of faults in photovoltaic arrays: a review. In: 2018 IEEE 12th international conference on compatibility, power electronics and power engineering (CPE-POWERENG 2018); 2018. p. 1–8. https://doi.org/10.1109/CPE.2018.8372609.
[19] Ahmed R, Sreeram V, Mishra Y, Arif MD. A review and evaluation of the state-of-the-art in PV solar power forecasting: techniques and optimization. Renew Sustain Energy Rev 2020;124:109792. https://doi.org/10.1016/j.rser.2020.109792.
[20] Rajagukguk RA, Ramadhan RAA, Lee H-J. A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. Energies 2020;13:6623. https://doi.org/10.3390/en13246623.
[21] Garud KS, Jayaraj S, Lee M-Y. A review on modeling of solar photovoltaic systems using artificial neural networks, fuzzy logic, genetic algorithm and hybrid models. Int J Energy Res 2021;45:6–35. https://doi.org/10.1002/er.5608.
[22] Ridha HM, Gomes C, Hizam H, Ahmadipour M, Heidari AA, Chen H. Multi-objective optimization and multi-criteria decision-making methods for optimal design of standalone photovoltaic system: a comprehensive review. Renew Sustain Energy Rev 2021;135:110202. https://doi.org/10.1016/j.rser.2020.110202.
[23] Li B, Delpha C, Diallo D, Migan-Dubois A. Application of artificial neural networks to photovoltaic fault detection and diagnosis: a review. Renew Sustain Energy Rev 2021;138:110512. https://doi.org/10.1016/j.rser.2020.110512.
[24] Navid Q, Hassan A, Fardoun AA, Ramzan R, Alraeesi A. Fault diagnostic methodologies for utility-scale photovoltaic power plants: a state of the art review. Sustainability 2021;13:1629. https://doi.org/10.3390/su13041629.
[25] Mellit A, Kalogirou S. Artificial intelligence and internet of things to improve efficacy of diagnosis and remote sensing of solar photovoltaic systems: challenges, recommendations and future directions. Renew Sustain Energy Rev 2021;143:110889. https://doi.org/10.1016/j.rser.2021.110889.
[26] Al-Mashhadani R, Alkawsi G, Baashar Y, Alkahtani AA, Nordin FH, Hashim W. Deep learning methods for solar fault detection and classification: a review. Information Sciences Letters 2021;10.
[27] Berghout T, Benbouzid M, Ma X, Djurović S, Mouss L-H. Machine learning for photovoltaic systems condition monitoring: a review. In: Iecon 2021 – 47th annual conference of the IEEE industrial electronics society; 2021. p. 1–5. https://doi.org/10.1109/IECON48115.2021.9589423.
[28] Hong Y-Y, Pula RA. Methods of photovoltaic fault detection and classification: a review. Energy Rep 2022;8:5898–929. https://doi.org/10.1016/j.egyr.2022.04.043.
[29] Waqar Akram M, Li G, Jin Y, Chen X. Failures of photovoltaic modules and their detection: a review. Appl Energy 2022;313:118822. https://doi.org/10.1016/j.apenergy.2022.118822.
[30] Gaviria JF, Narváez G, Guillen C, Giraldo LF, Bressan M. Machine learning in photovoltaic systems: a review. Renew Energy 2022;196:298–318. https://doi.org/10.1016/j.renene.2022.06.105.
[31] Ranjan S, Balaji S, Panella RA, Ydstie BE. Silicon solar cell production. Comput Chem Eng 2011;35:1439–53. https://doi.org/10.1016/j.compchemeng.2011.04.017.
[32] Buratti Y, Eijkens C, Hameiri Z. Optimization of solar cell production lines using neural networks and genetic algorithms. ACS Appl Energy Mater 2020;3:10317–22.
[33] Wagner-Mohnsen H, Esefelder S, Klöter B, Mitchell B, Schinke C, Bredemeier D, et al. Combining numerical simulations, machine learning and genetic algorithms for optimizing a POCl3 diffusion process. In: 2021 IEEE 48th photovoltaic specialists conference (PVSC); 2021. p. 528–31.
[34] Kayabasi E, Ozturk S, Celik E, Kurt H. Determination of cutting parameters for silicon wafer with a diamond wire saw using an artificial neural network. Sol Energy 2017;149:285–93. https://doi.org/10.1016/j.solener.2017.04.022.
[35] Kayabasi E, Ozturk S, Celik E, Kurt H, Arcaklioğlu E. Prediction of nano etching parameters of silicon wafer for a better energy absorption with the aid of an artificial neural network. Sol Energy Mater Sol Cell 2018;188:234–40. https://doi.org/10.1016/j.solmat.2018.08.027.

[36] Ozturk S, Kayabasi E, Celik E, Kurt H. Determination of lapping parameters for silicon wafer using an artificial neural network. J Mater Sci Mater Electron 2018; 29:260–70. https://doi.org/10.1007/s10854-017-7912-4.

[37] Ozturk S, Aydin L, Kucukdogan N, Celik E. Optimization of lapping processes of silicon wafer for photovoltaic applications. Sol Energy 2018;164:1–11. https://doi.org/10.1016/j.solener.2018.02.039.

[38] Haug H, Greulich J. PC1Dmod 6.2–Improved simulation of c-Si devices with updates on device physics and user interface. Energy Proc 2016;92:60–8.

[39] Rubinstein RY, Kroese DP. Simulation and the Monte Carlo method. second ed. 2007.

[40] Qi X, Ma W, Dang Y, Su W, Liu L. Optimization of the melt/crystal interface shape and oxygen concentration during the Czochralski silicon crystal growth process using an artificial neural network and a genetic algorithm. J Cryst Growth 2020; 548:125828. https://doi.org/10.1016/j.jcrysgro.2020.125828.

[41] Rachdi L, Hofmann M. Use of optical emission spectroscopy to predict silicon nitride layer properties. Vacuum 2021;191:110322. https://doi.org/10.1016/j.vacuum.2021.110322.

[42] Holland JH. Genetic algorithms and adaptation. In: Selfridge OG, Rissland EL, Arbib MA, editors. Adaptive control of ill-defined systems. Boston, MA: Springer US; 1984. p. 317–33. https://doi.org/10.1007/978-1-4684-8941-5_21.

[43] Aiello G, Alfonzetti S, Rizzo SA, Salerno N. Multi-objective optimization of thin-film silicon solar cells with metallic and dielectric nanoparticles. Energies 2017; 10:53. https://doi.org/10.3390/en10010053.

[44] Kaya M, Hajimirza S. Rapid optimization of external quantum efficiency of thin film solar cells using surrogate modeling of absorptivity. Sci Rep 2018;8:8170. https://doi.org/10.1038/s41598-018-26469-3.

[45] Hamedi S, Kordrostami Z, Yadollahi A. Artificial neural network approaches for modeling absorption spectrum of nanowire solar cells. Neural Comput Appl 2019; 31:8985–95. https://doi.org/10.1007/s00521-019-04406-3.

[46] Buratti Y, Le Gia QT, Dick J, Zhu Y, Hameiri Z. Extracting bulk defect parameters in silicon wafers using machine learning models. npj Comput Mater 2020;6:1–8. https://doi.org/10.1038/s41524-020-00410-7.

[47] Shockley W, Read WT. Statistics of the recombinations of holes and electrons. Phys Rev 1952;87:835–42. https://doi.org/10.1103/PhysRev.87.835.

[48] Hall RN. Electron-hole recombination in germanium. Phys Rev 1952;87:387. https://doi.org/10.1103/PhysRev.87.387. 387.

[49] Abdullah-Vetter Z, Dwivedi P, Buratti Y, Sowmya A, Trupke T, Hameiri Z. Advanced analysis of internal quantum efficiency measurements using machine learning. Prog Photovoltaics Res Appl 2023. https://doi.org/10.1002/pip.3683.

[50] Fischer B. Loss analysis of crystalline silicon solar cells using photoconductance and quantum efficiency measurements. Konstanz University; 2003. PhD Thesis.

[51] Zhu H, Yan W, Liu Y, Hu D, Tu Y, Huang Z, et al. Design investigation on 100 μm-thickness thin silicon PERC solar cells with assistance of machine learning. Mater Sci Semicond Process 2022;137:106198. https://doi.org/10.1016/j.mssp.2021.106198.

[52] Kumar V, Maheshwari P. Advanced analytics on IV curves and electroluminescence images of photovoltaic modules using machine learning algorithms. Prog Photovoltaics Res Appl 2021;30:880–8. https://doi.org/10.1002/pip.3469.

[53] Grau-Luque E, Guc M, Becerril-Romero I, Izquierdo-Roca V, Pérez-Rodríguez A, Bolt P, et al. Thickness evaluation of AlOx barrier layers for encapsulation of flexible PV modules in industrial environments by normal reflectance and machine learning. Prog Photovoltaics Res Appl 2021;30:229–39. https://doi.org/10.1002/pip.3478.

[54] Kato H, Kamibeppu S, Kojima T, Matsumoto T, Kudo H, Takeuchi Y, et al. Estimation of crystal orientation of grains on polycrystalline silicon substrate by recurrent neural network. IEEJ Trans Electr Electron Eng 2022;17:1685–7. https://doi.org/10.1002/tee.23676.

[55] Xie T, Wan Y, Wang H, Østrøm I, Wang S, He M, et al. Opinion mining by convolutional neural networks for maximizing discoverability of nanomaterials. J Chem Inf Model 2024;64:2746–59. https://doi.org/10.1021/acs.jcim.3c00746.

[56] Xie T, Wan Y, Huang W, Zhou Y, Liu Y, Linghu Q, et al. Large language models as master key: unlocking the secrets of materials science with GPT. https://doi.org/10.48550/arXiv.2304.02213; 2023.

[57] Ye Y, Ren J, Wang S, Wan Y, Razzak I, Xie T, et al. Construction of functional materials knowledge graph in multidisciplinary materials science via large language model. https://doi.org/10.48550/arXiv.2404.03080; 2024.

[58] Ma X, Huang W-H, Schnabel E, Köhl M, Brynjarsdóttir J, Braid JL, et al. Data-driven IV feature extraction for photovoltaic modules. IEEE J Photovoltaics 2019; 9:1405–12. https://doi.org/10.1109/JPHOTOV.2019.2928477.

[59] Liu J, Curran AJ, Fada JS, Ma X, Huang W-H, Jones CB, et al. Cross-correlation analysis of the indoor accelerated and real world exposed photovoltaic systems across multiple climate zones. In: 2018 IEEE 7th world conference on photovoltaic energy conversion (WCPEC) (A joint conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC); 2018. p. 3949–54. https://doi.org/10.1109/PVSC.2018.8547840.

[60] Wagner-Mohnsen H, Altermatt PP. A combined numerical modeling and machine learning approach for optimization of mass-produced industrial solar cells. IEEE J Photovoltaics 2020;10:1441–7.

[61] Olikh O, Lozitsky O, Zavhorodnii O. Estimation for iron contamination in Si solar cell by ideality factor: deep neural network approach. Prog Photovoltaics Res Appl 2022;30:648–60. https://doi.org/10.1002/pip.3539.

[62] Bucciarelli LL. Power loss in photovoltaic arrays due to mismatch in cell characteristics. Sol Energy 1979;23:277–88. https://doi.org/10.1016/0038-092X(79)90121-X.

[63] Wasmer S, Greulich J, Höffler H, Wöhrle N, Demant M, Fertig F, et al. Impact of material and process variations on the distribution of multicrystalline silicon PERC cell efficiencies. IEEE J Photovoltaics 2017;7:118–28. https://doi.org/10.1109/JPHOTOV.2016.2626145.

[64] Klöter B. Application of machine learning for production optimization. In: 2018 IEEE 7th world conference on photovoltaic energy conversion (WCPEC) (A joint conference of 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC); 2018. p. 3489–91. https://doi.org/10.1109/PVSC.2018.8547467.

[65] Evans R, Boreland M. Multivariate data analytics in PV manufacturing—four case studies using manufacturing datasets. IEEE J Photovoltaics 2018;8:38–47. https://doi.org/10.1109/JPHOTOV.2017.2778571.

[66] Arena E, Corsini A, Ferulano R, Iuvara DA, Miele ES, Ricciardi Celsi L, et al. Anomaly detection in photovoltaic production factories via Monte Carlo pre-processed principal component analysis. Energies 2021;14:3951. https://doi.org/10.3390/en14133951.

[67] Pei F-Q, Tong Y-F, Yuan M-H, Ding K, Chen X-H. The digital twin of the quality monitoring and control in the series solar cell production line. J Manuf Syst 2021; 59:127–37. https://doi.org/10.1016/j.jmsy.2021.02.001.

[68] Wasmer S, Hübener K, Klöter B. Explaining the efficiencies of mass-produced p-type Cz-si solar cells by interpretable machine learning. Sol RRL 2021;6:2100477. https://doi.org/10.1002/solr.202100477.

[69] Hossain MJ, Gregory G, Schneller EJ, Gabor AM, Blum AL, Yang Z, et al. A comprehensive methodology to evaluate losses and process variations in silicon solar cell manufacturing. IEEE J Photovoltaics 2019;9:1350–9. https://doi.org/10.1109/JPHOTOV.2019.2926628.

[70] Chan H-P, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis. Adv Exp Med Biol 2020;1213:3–21. https://doi.org/10.1007/978-3-030-33128-3_1.

[71] Wang Y, Li L, Sun Y, Xu J, Jia Y, Hong J, et al. Adaptive automatic solar cell defect detection and classification based on absolute electroluminescence imaging. Energy 2021;229:120606. https://doi.org/10.1016/j.energy.2021.120606.

[72] Meng Z, Xu S, Wang L, Gong Y, Zhang X, Zhao Y. Defect object detection algorithm for electroluminescence image defects of photovoltaic modules based on deep learning. Energy Sci Eng 2022;10:800–13. https://doi.org/10.1002/ese3.1056.

[73] Fioresi J, Colvin DJ, Frota R, Gupta R, Li M, Seigneur HP, et al. Automated defect detection and localization in photovoltaic cells using semantic segmentation of electroluminescence images. IEEE J Photovoltaics 2022;12:53–61. https://doi.org/10.1109/JPHOTOV.2021.3131059.

[74] Parikh HR, Buratti Y, Spataru S, Villebro F, Reis Benatto GAD, Poulsen PB, et al. Solar cell cracks and finger failure detection using statistical parameters of electroluminescence images and machine learning. Appl Sci 2020;10:8834. https://doi.org/10.3390/app10248834.

[75] Zhang X, Sun H, Zhou Y, Xi J, Li M. A novel method for surface defect detection of photovoltaic module based on independent component analysis. Math Probl Eng 2013;2013:e520568. https://doi.org/10.1155/2013/520568.

[76] Su B, Chen H, Zhu Y, Liu W, Liu K. Classification of manufacturing defects in multicrystalline solar cells with novel feature descriptor. IEEE Trans Instrum Meas 2019;68:4675–88. https://doi.org/10.1109/TIM.2019.2900961.

[77] Serfa Juan RO, Kim J. Photovoltaic cell defect detection model based-on extracted electroluminescence Images using SVM classifier. In: 2020 international conference on artificial intelligence in information and communication (ICAIIC); 2020. p. 578–82. https://doi.org/10.1109/ICAIIC48513.2020.9065065.

[78] Buerhop-Lutz C, Deitsch S, Maier A, Gallwitz F, Berger S, Doll B, et al. A benchmark for visual identification of defective solar cells in electroluminescence imagery. In: 35th European photovoltaic solar energy conference and exhibition; 2018. p. 1287–9.

[79] Deitsch S, Christlein V, Berger S, Buerhop-Lutz C, Maier A, Gallwitz F, et al. Automatic classification of defective photovoltaic module cells in electroluminescence images. Sol Energy 2019;185:455–68. https://doi.org/10.1016/j.solener.2019.02.067.

[80] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. In: 3rd International conference on learning representations; 2014. p. 1–14.

[81] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97. https://doi.org/10.1007/BF00994018.

[82] Akram MW, Li G, Jin Y, Chen X, Zhu C, Zhao X, et al. CNN based automatic detection of photovoltaic cell defects in electroluminescence images. Energy 2019;189:116319. https://doi.org/10.1016/j.energy.2019.116319.

[83] Acharya AK, Sahu PK, Jena SR. Deep neural network based approach for detection of defective solar cell. Mater Today Proc 2020;39:2009–14.

[84] Koch G. Siamese neural networks for one-shot image recognition n.d.

[85] Ge C, Liu Z, Fang L, Ling H, Zhang A, Yin C. A hybrid fuzzy convolutional neural network based mechanism for photovoltaic cell defect detection with electroluminescence images. IEEE Trans Parallel Distr Syst 2021;32:1653–64.

[86] Demirci MY, Beşli N, Gümüşçü A. Efficient deep feature extraction and classification for identifying defective photovoltaic module cells in electroluminescence images. Expert Syst Appl 2021;175:114810.

[87] Huang C, Zhang Z, Wang L. PSOPruner: PSO-based deep convolutional neural network pruning method for PV module defects classification. IEEE J Photovoltaics 2022:1–9.

[88] 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); 2018. p. 2035–9. https://doi.org/10.23919/EUSIPCO.2018.8553025.

[89] Ying Z, Li M, Tong W, Haiyong C. Automatic detection of photovoltaic module cells using multi-channel convolutional neural network. 2018 Chinese Automation Congress (CAC); 2018. p. 3571–6.
[90] Karimi AM, Fada JS, Hossain MA, Yang S, Peshek TJ, Braid JL, et al. Automated pipeline for photovoltaic module electroluminescence image processing and degradation feature classification. IEEE J Photovoltaics 2019;9:1324–35. https://doi.org/10.1109/JPHOTOV.2019.2920732.
[91] Luo Z, Cheng SY, Zheng QY. GAN-based augmentation for improving CNN performance of classification of defective photovoltaic module cells in electroluminescence images. IOP Conf Ser Earth Environ Sci 2019;354:012106. https://doi.org/10.1088/1755-1315/354/1/012106.
[92] 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. https://doi.org/10.1016/j.solener.2020.03.049.
[93] Verma S, Taluja HK, Chaudhary P. Automatic defect classification of electro-luminescence Images of photovoltaic modules based on deep learning CNN. Int J Mech Eng 2021;6:10.
[94] Et-taleby A, Chaibi Y, Allouhi A, Boussetta M, Benslimane M. A combined convolutional neural network model and support vector machine technique for fault detection and classification based on electroluminescence images of photovoltaic modules. Sustainable Energy, Grids and Networks 2022;32:100946. https://doi.org/10.1016/j.segan.2022.100946.
[95] Girshick R. Fast R-CNN. In: 2015 IEEE international conference on computer vision (ICCV); 2015. p. 1440–8. https://doi.org/10.1109/ICCV.2015.169.
[96] Shaoqing R, Kaiming H, Girshick R, Jian S. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 2017;39:1137–49. https://doi.org/10.1109/TPAMI.2016.2577031.
[97] Liu L, Zhu Y, Ur Rahman MR, Zhao P, Chen H. Surface defect detection of solar cells based on feature pyramid network and GA-Faster-RCNN. In: 2019 2nd China symposium on cognitive computing and hybrid intelligence (CCHI); 2019. p. 292–7.
[98] 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. https://doi.org/10.1016/j.infrared.2020.103334.
[99] Xu Z, Wu Z, Fan W. Improved SSD-assisted algorithm for surface defect detection of electromagnetic luminescence. Proc Inst Mech Eng O J Risk Reliab 2021;235 (5):761–8. https://doi.org/10.1177/1748006X21995388.
[100] Su B, Chen H, Zhou Z. BAF-Detector: an efficient CNN-based detector for photovoltaic cell defect detection. IEEE Trans Ind Electron 2021;69:3161–71.
[101] Su B, Chen H yong, Chen P, Bian G-B, Liu kun, Liu W. Deep learning-based solar-cell manufacturing defect detection with complementary attention network. IEEE Trans Ind Inf 2021;17:4084–95.
[102] Li L, Wang Z, Zhang T. GBH-YOLOv5: ghost convolution with bottleneckCSP and tiny target prediction head incorporating YOLOv5 for PV panel defect detection. Electronics 2023;12:561. https://doi.org/10.3390/electronics12030561.
[103] Huang J, Zeng K, Zhang Z, Zhong W. Solar panel defect detection design based on YOLO v5 algorithm. Heliyon 2023;9:e18826. https://doi.org/10.1016/j.heliyon.2023.e18826.
[104] Mayr M, Hoffmann M, Maier A, Christlein V. Weakly supervised segmentation of cracks on solar cells using normalized Lp norm. In: 2019 IEEE international conference on image processing (ICIP); 2019. p. 1885–9. https://doi.org/10.1109/ICIP.2019.8803116.
[105] Rahman MRU, Chen H, Xi W. U-Net based defects inspection in photovoltaic electroluminecscence images. In: 2019 IEEE international conference on big knowledge (ICBK); 2019. p. 215–20. https://doi.org/10.1109/ICBK.2019.00036.
[106] Jiang Y, Zhao C, Ding W, Hong L, Shen Q. Attention M-net for automatic pixel-level micro-crack detection of photovoltaic module cells in electroluminescence images. In: 2020 IEEE 9th data driven control and learning systems conference (DDCLS); 2020. p. 1415–21. https://doi.org/10.1109/DDCLS49620.2020.9275068.
[107] 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: The Official Journal of the International Neural Network Society 2020;127:132–40. https://doi.org/10.1016/j.neunet.2020.04.012.
[108] Otamendi U, Martinez I, Quartulli M, Olaizola IG, Viles E, Cambarau W. Segmentation of cell-level anomalies in electroluminescence images of photovoltaic modules. Sol Energy 2021;220:914–26. https://doi.org/10.1016/j.solener.2021.03.058.
[109] Pratt L, Govender D, Klein R. Defect detection and quantification in electroluminescence images of solar PV modules using U-net semantic segmentation. Renew Energy 2021;178:1211–22.
[110] Tian S, Li W, Li S, Tian G, Sun L, Ning X. Image defect detection and segmentation algorithm of solar cell based on convolutional neural network. In: 2021 6th international conference on intelligent computing and signal processing (ICSP); 2021. p. 154–7.
[111] Jiang Y, Zhao C. Attention classification-and-segmentation network for micro-crack anomaly detection of photovoltaic module cells. Sol Energy 2022;238: 291–304.
[112] Wang C, Chen H, Zhao S, Rahman MRU. Efficient and refined deep convolutional features network for the crack segmentation of solar cell electroluminescence images. IEEE Trans Semicond Manuf 2022;35:610–9.
[113] Ni B, Zou P, Li Q, Chen Y. Intelligent defect detection method of photovoltaic modules based on deep learning. Advances in Intelligent Systems Research 2018; 161:167–73. https://doi.org/10.2991/tlicsc-18.2018.27.
[114] Buratti Y, Sowmya A, Dumbrell R, Dwivedi P, Trupke T, Hameiri Z. Automated efficiency loss analysis by luminescence image reconstruction using generative adversarial networks. Joule 2022;6:1320–32. https://doi.org/10.1016/j.joule.2022.05.001.
[115] Su B, Zhou Z, Chen H, Cao X. SIGAN: a novel image generation method for solar cell defect segmentation and augmentation. ArXiv 2021:e2104.04953.
[116] Balzategui J, Eciolaza L, Maestro-Watson D. Anomaly detection and automatic labeling for solar cell quality inspection based on generative adversarial network. Sensors 2021;21:4361.
[117] Buratti Y, Sowmya A, Evans R, Trupke T, Hameiri Z. Half and full solar cell efficiency binning by deep learning on electroluminescence images. Prog Photovoltaics Res Appl 2021;30:276–87.
[118] Kunze P, Greulich JM, Tummalieh A, Wirtz W, Hoeffler H, Woehrle N, et al. Contactless inline IV measurement of solar cells using an empirical model. Sol RRL 2022;7:2200599. https://doi.org/10.1002/solr.202200599.
[119] Dwivedi P, Weber JW, Lee Chin R, Trupke T, Hameiri Z. Deep learning method for enhancing luminescence image resolution. Sol Energy Mater Sol Cell 2023;257: 112357. https://doi.org/10.1016/j.solmat.2023.112357.
[120] Liu G, Dwivedi P, Trupke T, Hameiri Z. Deep learning model to denoise luminescence images of silicon solar cells. Adv Sci 2023:2300206. https://doi.org/10.1002/advs.202300206.
[121] Demant M, Virtue P, Kovvali A, Yu SX, Rein S. Learning quality rating of as-cut mc-Si wafers via convolutional regression networks. IEEE J Photovoltaics 2019;9: 1064–72.
[122] Demant M, Virtue P, Kovvali A, Yu SX, Rein S. Visualizing material quality and similarity of mc-Si wafers learned by convolutional regression networks. IEEE J Photovoltaics 2019;9:1073–80. https://doi.org/10.1109/JPHOTOV.2019.2906037.
[123] Demant M, Kurumundayil L, Kunze P, Woernhoer A, Kovvali A, Rein S. Machine learning for advanced solar cell production: adversarial denoising, sub-pixel alignment and the digital twin. In: NeurIPS 2020 workshop tackling climate change with machine learning; 2020. p. 7.
[124] Kurumundayil L, Schremmer H, Rein S, Demant M. Denoising of image data for DWS wafer characterization using generative adversarial networks. In: 38th European photovoltaic solar energy conference and exhibition; 2021. p. 194–7. https://doi.org/10.4229/EUPVSEC20212021-2CV.1.10.
[125] Fu Y, Li X, Ma X. Deep-learning-based defect evaluation of mono-like cast silicon wafers. Photonics 2021;8:426. https://doi.org/10.3390/photonics8100426.
[126] Han H, Gao C, Zhao Y, Liao S, Tang L, Li X. Polycrystalline silicon wafer defect segmentation based on deep convolutional neural networks. Pattern Recogn Lett 2020;130:234–41. https://doi.org/10.1016/j.patrec.2018.12.013.
[127] Bhattacharya S, Arya D, Bhowmick D, Thomas RM, Gupta DK. Improving solar cell metallization designs using convolutional neural networks. ArXiv 2021;: e2104.04017.
[128] Karimi AM, Fada JS, Parrilla NA, Pierce BG, Koyutürk M, French RH, et al. Generalized and mechanistic PV module performance prediction from computer vision and machine learning on electroluminescence images. IEEE J Photovoltaics 2020;10:878–87. https://doi.org/10.1109/JPHOTOV.2020.2973448.
[129] Hoffmann M, Buerhop-Lutz C, Reeb L, Pickel T, Winkler T, Doll B, et al. Deep-learning-based pipeline for module power prediction from electroluminescense measurements. Prog Photovoltaics Res Appl 2021;29:920–35. https://doi.org/10.1002/pip.3416.
[130] Kunwar A, Malla PB, Sun J, Qu L, Ma H. Convolutional neural network model for synchrotron radiation imaging datasets to automatically detect interfacial microstructure: an in situ process monitoring tool during solar PV ribbon fabrication. Sol Energy 2021;224:230–44. https://doi.org/10.1016/j.solener.2021.06.006.
[131] Li X, Li W, Yang Q, Yan W, Zomaya AY. An unmanned inspection system for multiple defects detection in photovoltaic plants. IEEE J Photovoltaics 2020;10: 568–76. https://doi.org/10.1109/JPHOTOV.2019.2955183.
[132] Chen H, Pang Y, Hu Q, Liu K. Solar cell surface defect inspection based on multispectral convolutional neural network. J Intell Manuf 2020;31:453–68. https://doi.org/10.1007/s10845-018-1458-z.
[133] Pierdicca R, Malinverni ES, Piccinini F, Paolanti M, Felicetti A, Zingaretti P. Deep convolutional neural network for automatic detection of damaged photovoltaic cells. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2018;XLII–2:893–900. https://doi.org/10.5194/isprs-archives-XLII-2-893-2018.
[134] 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. https://doi.org/10.1016/j.solener.2020.01.055.
[135] Bommes L, Pickel T, Buerhop-Lutz C, Hauch J, Brabec C, Peters IM. Computer vision tool for detection, mapping, and fault classification of photovoltaics modules in aerial IR videos. Prog Photovoltaics Res Appl 2021;29:1236–51. https://doi.org/10.1002/pip.3448.
[136] Fonseca Alves RH, Antero de Deus Júnior G, Marra EG, Lemos RP. Automatic fault classification in photovoltaic modules using convolutional neural networks. Renew Energy 2021;179:502–16. https://doi.org/10.1016/j.renene.2021.07.070.
[137] Zefri Y, Sebari I, Hajji H, Aniba G. Developing a deep learning-based layer-3 solution for thermal infrared large-scale photovoltaic module inspection from orthorectified big UAV imagery data. Int J Appl Earth Obs Geoinf 2022;106: 102652. https://doi.org/10.1016/j.jag.2021.102652.
[138] Su B, Chen H, Liu K, Liu W. RCAG-Net: residual channelwise attention gate network for hot spot defect detection of photovoltaic farms. IEEE Trans Instrum Meas 2021;70:1–14. https://doi.org/10.1109/TIM.2021.3054415.
[139] Gilleland B, Hobbs WB, Richardson JB. High throughput detection of cracks and other faults in solar PV modules using a high-power ultraviolet fluorescence

imaging system. 46th IEEE Photovoltaic Specialist Conference 2019:2575–82. https://doi.org/10.1109/PVSC40753.2019.8981262.
[140] FAIR Principles. GO FAIR n.d. https://www.go-fair.org/fair-principles/. [Accessed 15 April 2024].
[141] French RH, Karimi AM, Braid JL. Electroluminescent (EL) image dataset of PV module under step-wise damp heat exposures. https://doi.org/10.17605/OSF.IO/4QRTV; 2019.
[142] Millendorf M, Obropta E, Vadhavkar N. Infrared solar module dataset for anomaly detection. In: International conference on learning representation; 2020. p. 1–5.
[143] Hoffman RR, Mueller ST, Klein G, Litman J. Metrics for explainable AI: challenges and prospects. https://doi.org/10.48550/arXiv.1812.04608; 2019.
[144] Meyes R, Lu M, de Puiseau CW, Meisen T. Ablation studies in artificial neural networks. https://doi.org/10.48550/arXiv.1901.08644; 2019.

Supporting documentation

This document contains another copy of the abbreviations and symbols used in the main review article. The next section also contains a background on machine learning concepts that may be useful for readers who may need additional background to understand the papers reviewed in the article.

Abbreviations

AB	Adaptive boosting
AE	Autoencoder
AI	Artificial intelligence
Al-BSF	Aluminium back surface field
AlO_x	Aluminium oxide
a-Si	Amorphous silicon
B	Boron
BAF	Bi-directional attention feature pyramid network
CAN	Complementary attention network
CIGS	Copper indium gallium selenide
CNN	Convolutional neural network
CV	Cross-validation
Cz-Si	Czochralski silicon
DBN	Deep belief network
DEA	Differential evolution algorithm
EL	Electroluminescence
EQE	External quantum efficiency
ERDCF	Efficient and refined deep convolutional features
FDTD	Finite difference time domain
Fe	Iron
FEM	Finite element method
FF	Fill factor
FN	False positive
FP	False negative
FPN	Feature pyramid network
GA	Genetic algorithm
GAB	Global attention block
GAN	Generative adversarial network
GAR	Guided anchoring region
GR	Gaussian regression
HJT	Heterojunction
IoU	Intersection over union
IPCC	Intergovernmental panel on climate change
IRT	Infrared thermography
ITRPV	International roadmap for photovoltaic

IQE	Internal quantum efficiency
I-V	Current-voltage
k-NN	k-nearest neighbours
LDA	Linear discriminant analysis
LSTM	Long short-term memory
MAE	Mean absolute error
mAP	Mean average precision
mc-Si	Multicrystalline silicon
ML	Machine learning
mono-Si	Monocrystalline silicon
NMA	Nelder-Mead algorithm
NN	Neural network
OES	Optical emissions spectroscopy
PCA	Principal component analysis
PECVD	Plasma-enhanced chemical vapor deposition
PERC	Passivated emitter and rear contact
PL	Photoluminescence
$POCl_3$	Phosphorus oxychloride
PSALHE-EA	Parallel self-adaptive low-high evaluation evolutionary algorithm
PSO	Particle swarm optimisation
PV	Photovoltaic
R-CNN	Region-proposal convolutional neural network
RF	Random forest
RMSE	Root mean square error
RSA	Random search algorithm
SAA	Simulated annealing algorithm
S-ALD	Spatial atomic layer deposition
SE	Squeeze and excitation blocks
SEM	Structural equation modelling
Si	Silicon
SiN_x	Silicon nitride
SiO_x	Silicon oxide
SRH	Shockley Read Hall
SSD	Single-shot detection
SSIM	Structure similarity index
SVM	Support vector machine
TN	True negative
TP	True positive
UVF	Ultra-violet fluorescence
YOLO	You only look once

Symbols

$b_{i,c}$	Binary indicator of element *i* in class *c*
F_1	Harmonic mean of precision and recall
I_{sc}	Short-circuit current [A]
J_{sc}	Short-circuit current density [A/cm²]
M	Number of classes
N	Dataset size
$p_{i,c}$	Predicted probability of element *i* in class *c*
P_{mp}	Maximum power point power [W]
R^2	Coefficient of determination
RE_{max}	Maximum relative error
R_s	Series resistance [Ω]
R_{sh}	Shunt resistance [Ω]
V_{oc}	Open-circuit voltage [V]
y_i	i-th value
y_{mean}	Mean value
y^{pred}	Predicted value
y^{true}	Actual value
σ_n	Electron capture cross-section [cm²]
σ_p	Hole capture cross-section [cm²]
E_t	Defect energy level [eV]
k	Capture cross-section ratio

A. Machine learning concepts

The field of artificial intelligence (AI) encompasses various subfields, among which machine learning (ML) stands out for its ability to learn and improve without explicit programming. By using ML training algorithms, data-driven models are produced that can identify patterns, make predictions, and inform decision-making. With the power of ML, companies can streamline their operations, make more accurate forecasts, and develop innovative solutions to complex problems.

A.1. Classification and regression problems

ML algorithms can be broadly classified into three categories, each suited to different types of data and desired outcomes. The first category is *supervised learning*, where models make predictions based on labelled examples [1]. The second is *unsupervised learning*, which involves finding hidden patterns in unlabelled data [2]. The third category is *reinforcement learning*, where models learn by trial and error on specific tasks [3]. This review will mainly focus on supervised learning, which can be further subdivided into classification and regression tasks [4]. In classification tasks, the aim is to predict the label of an input from a discrete set of labels, whereas regression tasks involve predicting a continuous variable [4]. Throughout this section, the different types of ML algorithms, as well as their general training and evaluation frameworks are explored.

A.1.1. General pipeline

ML algorithms require a specific training process in which the dataset is divided into two sets: a training dataset usually consisting of 90% of the data, and a validation dataset consisting of the remaining 10%, as illustrated in Fig. 1(a). Both datasets contain feature vectors (inputs) and their corresponding labels (predictions). During the training phase, the ML algorithm uses the input of the training set to predict the output, which is then compared to the actual output to compute the error. This error is then used to update the weights associated with each feature, allowing the algorithm to improve its predictions. Once trained, the ML model is evaluated on a previously unseen validation dataset using appropriate metrics (which will be discussed later in this section). Evaluating the model on a separate, unseen dataset helps to avoid data leakage, which could artificially improve the results. While the training/validation ratio can be decided by the user, it is essential to ensure that the validation dataset is of a statistically significant size.

As the training and validation splitting is random, different splits can lead to different evaluation scores, presenting a limitation in the evaluation process. To mitigate this, cross-validation (CV) [5] is commonly employed, as it helps to yield a more robust evaluation of the model's performance. The principle of CV involves splitting the dataset into a fixed number of folds, where for each iteration of the CV, one of the folds is used for testing (like the validation dataset), while the remaining folds are used to train the ML algorithm. An example of a five-fold CV is illustrated in Fig. 1(b), where a validation set

(grey block) is still held out from the cross-validation for the final evaluation of the algorithm. When comparing ML algorithms for the same task, a CV approach (or in general multiple random train/validation splits) is preferred as it includes the random aspect of the splitting procedure [5]. If the CV approach is used for ML hyper-parameters (i.e., parameters controlling the training process of the ML model) or model selection, it is necessary to reserve a portion of the dataset for the final evaluation.

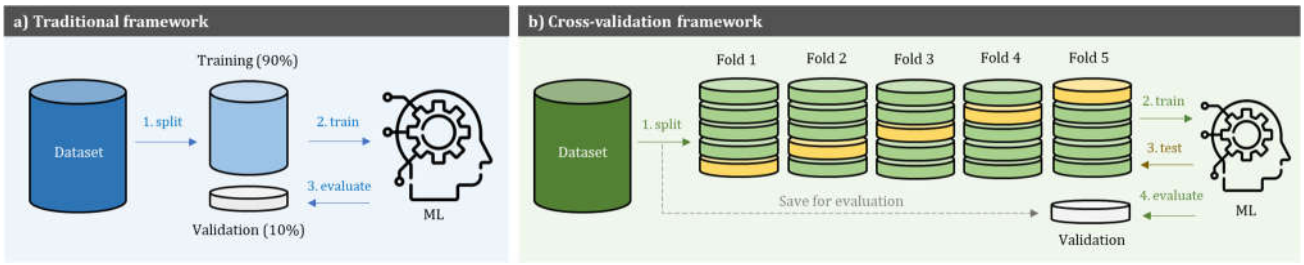


Figure 1 - *General framework for regression and classification problems using a (a) traditional and (b) cross-validation approach*

A.1.2. Common algorithms

There are various types of algorithms for regression and classification problems. Each is uniquely designed to address specific tasks. By selecting an appropriate ML algorithm that is suited to the specific task at hand, one can achieve improved accuracy and efficiency in the learning and prediction process. Below are some examples of traditional ML algorithms.

k-nearest neighbours (k-NN) [6] is a simple yet powerful ML approach that involves assigning the predicted label to be the majority vote of the closest known training data points. To build successful k-NN models, it is crucial to normalise the features and assign weights to the contribution of the neighbours. However, when dealing with skewed data distributions or over-represented subclasses, the k-NN algorithm may become biased and lose its ability to generalise to new data points [7]. Therefore, it is essential to carefully consider the data distribution and pre-processing steps before applying k-NN to a given problem [7].

Naïve-Bayes [8,9] is a powerful probabilistic approach to classification. It is based on the Bayes theorem, which calculates the conditional probability of a class given the observed features and prior probabilities. In Naïve Bayes, the outcome is the class with the highest conditional probability, calculated using the maximum likelihood approach. Despite its simplicity, Naïve Bayes can often achieve high accuracy on classification tasks and is particularly effective when dealing with high-dimensional data [10]. However, it assumes independence between the features, which may not hold true in some cases, leading to suboptimal performance [10].

Random forest (RF) [11] is a widely used ensemble learning method for classification and regression problems. The RF algorithm is constructed by combining multiple decision trees, where each tree is built using a random subset of the training data and features. In RF, the decision of the majority of trees is taken to make a final prediction. One of the advantages of RF is that it can handle large datasets with

high-dimensional features and noisy data [11]. The performance of RF can be further improved through boosting techniques such as adaptive boosting (AB) [12,13] and gradient boosting [14], which aim to improve the accuracy of the model by assigning higher weights to the misclassified data points in the training process. These techniques can help to mitigate the overfitting problem and enhance the predictive power of the RF model [12–14].

Support vector machine (SVM) [15,16] is also a popular algorithm used for classification and regression tasks. SVM works by mapping the training data to a higher-dimensional feature space where it becomes easily separable. This mapping function can use a non-linear kernel, allowing for more complex model behaviour. SVMs are effective for high dimensional data with low sample size, making them memory efficient [4]. SVMs are particularly useful when the decision boundary between classes is non-linear, and the algorithm can find the best hyperplane that separates the classes with the maximum margin [4].

Neural networks (NN) [17] are a class of ML models inspired by the structure and function of the human brain. They consist of layers of interconnected nodes, called neurons, with each neuron applying a mathematical operation to the input and passing the result to the next layer. The weights of the connections between the neurons are updated during training through forward and backward propagation, where the model learns to optimise its predictions by adjusting the weights. NNs can learn complex non-linear patterns, through the use of non-linear activation functions in each layer [4].

Deep learning algorithms have become increasingly common in recent years. These algorithms are developed by extending the structure of neural networks, creating models with multiple layers that can learn complex representations of data. Below are a few examples of deep learning algorithms for regression and classification problems.

Convolutional neural network (CNN) [18] is an extension of traditional neural networks that is specifically designed to work with two-dimensional data, such as images. Instead of using a simple feature vector, CNNs consist of multiple layers of convolutional and pooling layers that extract important features from the input data. These layers are trained using adjustable weights, which are optimised during training to minimise a loss function. The feature extraction block of the CNN transforms the input image into a one-dimensional feature vector, which is then fed into a fully connected neural network for classification or regression, depending on the task at hand. This modular structure enables CNNs to achieve exceptional accuracy on a wide range of computer vision tasks for object recognition [19], image segmentation [20] and more [21,22].

Generative adversarial networks (GAN) [23] combine two CNNs trained in competition with one another. The first CNN, known as the generator, is designed to create synthetic data, typically images, based on either random noise or some context input. The second CNN, called the discriminator, then attempts to distinguish between real and fake data, classifying the incoming data as either real or fake. During training, the generator and discriminator networks engage in a zero-sum game, with the

generator attempting to produce more realistic data to deceive the discriminator, while the discriminator attempts to distinguish between real and fake data more effectively. The goal is to train both networks until the generator can create synthetic data that is indistinguishable from the real data, while the discriminator can correctly identify both real and fake data.

Region-proposal CNN (R-CNN) [24] uses CNN feature extraction to identify regions of interest in an image, simplifying the detection process to a handful of relevant regions. These regions are then typically passed to a CNN classifier to identify and label the different features present in each region. To improve the training speed of this process, Fast R-CNN [25,26] was developed. Fast R-CNNs can significantly increase the training speed by compressing the proposed regions into a single feature map. This is done by only running the entire image through the CNN once, as opposed to using each region of interest as a separate input for the CNN [25]. These features can be fed directly into a fully connected NN. Faster R-CNNs further reduce the computational cost by utilising only a single CNN to both generate the regions of interest as well as extract the features from those regions (as opposed to separate steps that are used in standard CNNs) [26].

You only look once **(YOLO)** [27] is another detection algorithm that is based on CNNs. Unlike other object detection methods, YOLO processes the entire image only once and predicts the class probabilities and bounding boxes for objects in the image simultaneously. This parallel prediction approach makes YOLO faster and more efficient than other methods like R-CNN, which perform region proposals and classification in separate steps. Furthermore, YOLO's unified architecture enables it to reason globally about the image when making predictions, leading to fewer false positives and more accurate object detection [27].

U-Net [20] is a deep learning architecture that is widely used for image segmentation tasks, such as identifying and labelling specific objects within an image. It is based on a "fully convolutional network," and consists of two key parts: the contracting path and the expansive path. Instead of using pooling layers, the contracting path employs down-sampling (encoder) layers, while the expansive path uses up-sampling (decoder) layers. Additionally, the feature maps from the encoding stage are concatenated with the feature maps in the decoding stage. This allows the network to preserve spatial information and produce a segmentation map of the same size as the input image [20].

Other architectures have been developed due to the flexibility of designing deep learning networks, as CNNs contain millions of trainable parameters [18]. Researchers are constantly designing and optimising CNN architectures for specific tasks, with the annual ImageNet [28] challenge being a driving force for the development of highly efficient CNNs. Popular architectures such as AlexNet,[21] ResNet [29], SqueezeNet [30], and VGGNet [31] have emerged from this effort. These architectures are often pre-trained on ImageNet and their weights are made publicly available, allowing for transfer learning, or fine-tuning on specific datasets.

A.1.3. Evaluation metrics

To evaluate the ML and deep learning algorithms, different metrics are used depending on the learning task. For regression problems, where the predicted and actual labels are continuous variables, several metrics are commonly employed. This includes the coefficient of determination (R^2), the root mean square error (RMSE), and the mean absolute error (MAE), defined as [32]:

$$R^2 := 1 - \frac{\sum_{i=1}^{N}\left(y_{\mathrm{i}}^{\mathrm{pred}} - y_{\mathrm{i}}^{\mathrm{true}}\right)^2}{\sum_{i=1}^{N}\left(y_{\mathrm{i}}^{\mathrm{true}} - y_{\mathrm{mean}}^{\mathrm{true}}\right)^2}, \tag{1}$$

$$\mathrm{RMSE} := \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_{\mathrm{i}}^{\mathrm{pred}} - y_{\mathrm{i}}^{\mathrm{true}}\right)^2}, \tag{2}$$

$$\mathrm{MAE} := \frac{1}{N}\sum_{i=1}^{N}\left|y_{\mathrm{i}}^{\mathrm{pred}} - y_{\mathrm{i}}^{\mathrm{true}}\right|, \tag{3}$$

where y^{pred} and y^{true} denote the predicted and true labels, respectively. The parameter $y_{\mathrm{mean}}^{\mathrm{true}}$ denotes the average values of the true labels. N is the total number of data points in the dataset.

The R^2 score is a metric with a maximum value of unity indicating a perfect fit of the data. This score enables comparison between prediction targets, making it a useful tool in assessing model performance. The RMSE and MAE are two other commonly used metrics in regression analysis. Both of these metrics have the same unit as the prediction target and are proportional to 1-R^2 [32]. However, the RMSE penalises larger errors more than the MAE due to its quadratic nature. Another useful regression scoring technique is to report the maximum relative errors (RE_{max}) over the dataset. This gives an idea of the worst possible performance for a given model and validation set, but it does not provide any information about the average performance of the model on the validation set. It is therefore important to use a combination of these metrics to gain a comprehensive understanding of the performance of regression models.

For classification, the predicted and true labels are discrete and compared using a confusion matrix. For instance, in binary classification, there are four possible outcomes: true positive (TP), true negative (TN), false negative (FN), and false positive (FP). The main classification metrics are defined as [4]:

$$\mathrm{Accuracy} := \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}}, \tag{4}$$

$$\mathrm{Recall} := \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}, \tag{5}$$

$$\text{Precision} := \frac{\text{TP}}{\text{TP+FP}}, \tag{6}$$

$$F_1 := \frac{\text{Recall}\cdot\text{Precision}}{2\cdot(\text{Recall+Precision})} = \frac{2\cdot\text{TP}}{2\cdot\text{TP+FN+FP}}. \tag{7}$$

The accuracy measures the ratio between correctly predicted class labels and the total number of instances in the dataset. The F_1-score is the harmonic average of the precision (how many of all predictions are correct) and recall (how many correct predictions are obtained out of all the true instances). The F_1-score is preferred for evaluation as it inherently considers class imbalance in the dataset, contrary to the accuracy [33]. For multiclass classification tasks, the metrics can be averaged for each label, treating the problem as a binary classification task of the label versus all other labels.

When applying deep learning to images, detection algorithms use metrics such as intersection over union (IoU) [34], mean average precision (mAP), or the Dice coefficient [35] to measure the overlap between predicted and ground truth regions. For segmentation or denoising tasks, the structure similarity index (SSIM) [36] is employed to compare the luminance, contrast, and structural components of the images.

A.2. Optimisation problems

Optimisation problems require the identification of the optimal solution, often with the objective of maximising or minimising a specific objective function. These problems often include constraints that define the range of valid solutions by setting limits on the values of parameters or variables. In this context, ML-assisted optimisation techniques that can be applied to process optimisation in PV are presented.

A.2.1. General pipeline

ML-assisted process optimisation involves a series of sequential steps, as illustrated in Fig. 2. The first step is the generation of datasets through numerical simulations and/or experiments, which contain input-output pairs representing different input process parameters and their corresponding outputs. The availability of a large amount of data is preferred for this step. In the second step, regression models are trained using the generated datasets from Step 1. The goal of this step is to capture the underlying relationships between the inputs and outputs in the dataset. These models can then be used to generate additional data more quickly.

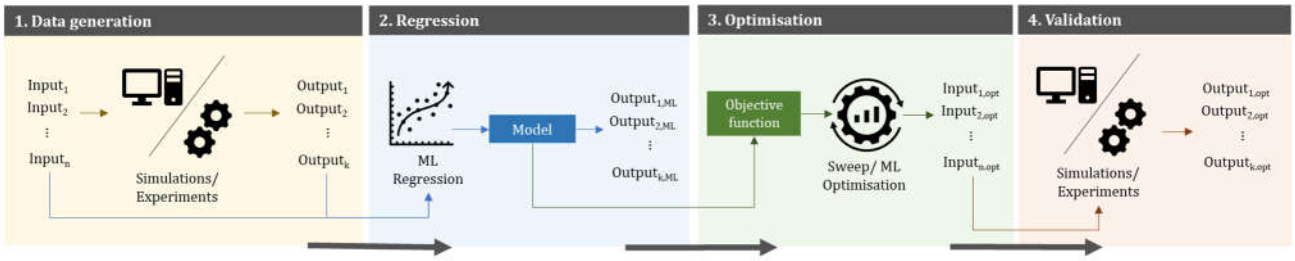


Figure 2 - *General framework for ML-assisted process optimisation*

Next, the regression models (obtained in Step 2) serve as the objective function for an optimisation problem that seeks to minimise or maximise the objective function while staying within a specific parameter space constraint. Various optimisation techniques can be used to identify the optimal input parameter values that would lead to the desired target output. Finally, the optimised input parameters from the previous step are validated using the simulated and/or experimental datasets generated in Step 1.

A.2.2. Common algorithms

There are various optimisation algorithms available, each with its unique strengths and weaknesses. The choice of algorithm often depends on the problem at hand, the type of data, and the computational resources available. Below are some examples of common optimisation algorithms.

Particle swarm optimisation (PSO) [37] is a popular optimisation algorithm that uses a swarm of particles to search for the optimal solution. Each particle is assigned a position and velocity in the parameter space, and the algorithm evaluates each particle at its position. The particle's position and velocity are then updated based on the local context, and the swarm iteratively converges towards multiple local optima. The global optimum is identified as the solution that contains the most particles. Its ability to efficiently search a large search space and find the global optimum makes it an attractive option for solving complex optimisation problems [38].

Genetic algorithm (GA) [39] is an optimisation algorithm inspired by Darwinian evolution. In GA, a population of randomly generated individuals is assigned a fitness score that can be evaluated using a data-driven model or an ML algorithm. The fittest individuals are selected to generate the next generation of individuals through breeding, crossover, and mutation operations. This process continues for multiple generations until the individuals in the population optimise the fitness score.

A.2.3. Evaluation metrics

Optimisation algorithms are typically assessed based on their convergence speed and the extent to which the calculated optimal inputs improve the initial output. The latter is typically evaluated by applying the optimised inputs to experiments and measuring the resulting output. Unlike regression and classification problems, there are no well-defined metrics for evaluating optimisation models. Consequently, the assessment of optimisation models tends to be more subjective and dependent on the

specific problem being addressed. Nonetheless, the practical implementation and performance of the optimised inputs in real-world applications remain a crucial aspect of the evaluation process.

References

[1] Russel S, Norvig P. Artificial intelligence: a modern approach. 4th ed. Pearson Education; 1995.

[2] Hinton G, Sejnowski TJ. Unsupervised learning: foundations of neural computation. MIT Press; 1999.

[3] Sutton RS, Barto AG. Reinforcement Learning, second edition: An Introduction. MIT Press; 2018.

[4] Raschka S, Mirjalili V. Python Machine Learning. 3rd ed. Birmingham - UK: Packt Publishing; 2019.

[5] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection 1995;2:1137–43.

[6] Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician 1992;46:175–85. https://doi.org/10.1080/00031305.1992.10475879.

[7] Kramer O. K-Nearest Neighbors. In: Kramer O, editor. Dimensionality Reduction with Unsupervised Nearest Neighbors, Berlin, Heidelberg: Springer; 2013, p. 13–23. https://doi.org/10.1007/978-3-642-38652-7_2.

[8] Stigler SM. Thomas Bayes's bayesian Inference. Journal of the Royal Statistical Society: Series A (General) 1982;145:250–8. https://doi.org/10.2307/2981538.

[9] Hand DJ, Yu K. Idiot's Bayes—not so stupid after all? International Statistical Review 2001;69:385–98. https://doi.org/10.1111/j.1751-5823.2001.tb00465.x.

[10] Raschka S. Naive Bayes and text classification I - introduction and theory. ArXiv 2014.

[11] Breiman L. Random forests. Machine Learning 2001;45:5–32. https://doi.org/10.1023/A:1010933404324.

[12] Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences 1997;55:119–39. https://doi.org/10.1006/jcss.1997.1504.

[13] Drucker H. Improving regressors using boosting techniques. ICML, 1997, p. 107–15.

[14] Friedman JH. Stochastic gradient boosting. Computational Statistics & Data Analysis 2002;38:367–78. https://doi.org/10.1016/S0167-9473(01)00065-2.

[15] Platt JC. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in Large Margin Classifiers, MIT Press; 1999, p. 61–74.

[16] Chang C-C, Lin C-J. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2011;2:1–27. https://doi.org/10.1145/1961189.1961199.

[17] Hinton GE. Connectionist learning procedures. Artificial Intelligence 1989;40:185–234. https://doi.org/10.1016/0004-3702(89)90049-0.

[18] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44. https://doi.org/10.1038/nature14539.

[19] Yadav SS, Jadhav SM. Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data 2019;6:113. https://doi.org/10.1186/s40537-019-0276-2.

[20] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham: Springer International Publishing; 2015, p. 234–41. https://doi.org/10.1007/978-3-319-24574-4_28.

[21] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25 2012:1097–105.

[22] Chan H-P, Samala RK, Hadjiiski LM, Zhou C. Deep learning in medical image analysis. Advances in Experimental Medicine and Biology 2020;1213:3–21. https://doi.org/10.1007/978-3-030-33128-3_1.

[23] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial networks. Advances in Neural Information Processing Systems, vol. 27, Curran Associates, Inc.; 2014.
[24] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, p. 580–7. https://doi.org/10.1109/CVPR.2014.81.
[25] Girshick R. Fast R-CNN. 2015 IEEE International Conference on Computer Vision (ICCV), 2015, p. 1440–8. https://doi.org/10.1109/ICCV.2015.169.
[26] Shaoqing R, Kaiming H, Girshick R, Jian S. Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 2017;39:1137–49. https://doi.org/10.1109/TPAMI.2016.2577031.
[27] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, p. 779–88. https://doi.org/10.1109/CVPR.2016.91.
[28] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet large scale visual recognition challenge. Int J Comput Vis 2015;115:211–52. https://doi.org/10.1007/s11263-015-0816-y.
[29] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA: IEEE; 2016, p. 770–8. https://doi.org/10.1109/CVPR.2016.90.
[30] Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. arXiv:160207360 [Cs] 2016.
[31] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv:14091556 [Cs] 2014.
[32] Glantz SA, Slinker BK. Primer of applied regression and analysis of variance. New York (N.Y.) : McGraw-Hill; 1990.
[33] Glunz SW, Rein S, Warta W, Knobloch J, Wettling W. Degradation of carrier lifetime in Cz silicon solar cells. Solar Energy Materials and Solar Cells 2001;65:219–29.
[34] Jaccard P. The distribution of the flora in the alpine zone.1. New Phytologist 1912;11:37–50. https://doi.org/10.1111/j.1469-8137.1912.tb05611.x.
[35] Dice LR. Measures of the smount of ecologic association between species. Ecology 1945;26:297–302. https://doi.org/10.2307/1932409.
[36] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 2004;13:600–12. https://doi.org/10.1109/TIP.2003.819861.
[37] Kennedy J, Eberhart R. Particle swarm optimization. Proceedings of ICNN'95 - International Conference on Neural Networks, vol. 4, 1995, p. 1942–8. https://doi.org/10.1109/ICNN.1995.488968.
[38] Gad AG. Particle swarm optimization algorithm and its applications: A systematic review. Archives of Computational Methods in Engineering 2022;29:2531–61. https://doi.org/10.1007/s11831-021-09694-4.
[39] Holland JH. Genetic algorithms and adaptation. In: Selfridge OG, Rissland EL, Arbib MA, editors. Adaptive Control of Ill-Defined Systems, Boston, MA: Springer US; 1984, p. 317–33. https://doi.org/10.1007/978-1-4684-8941-5_21.