Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/MBE

MBE, 23(5): 1269–1288.
DOI: 10.3934/mbe.2026046
Received: 12 June 2025
Revised: 14 January 2026
Accepted: 30 January 2026
Published: 01 April 2026

Research article

CNN-based estimation of series resistance in photovoltaic cells from electroluminescence images with application to output power prediction

Hector Felipe Mateo-Romero[1,5,*], Jose Ignacio Morales Aragonés[2], Luis Hernández-Callejo[3], Miguel Ángel González-Rebollo[4], Valentín Cardeñoso-Payo[5], Victor Alonso-Gómez[6], Mario Carbonó delaRosa[7] and Ginés García Mateos[8]

[1] Departamento Informática, Universidad Autónoma de Madrid, Madrid, 28049, Spain (until August 2025)
[2] Instituto de Nanociencia y Materiales de Aragón (INMA), CSIC-UNIZAR, Zaragoza, 50018, Spain
[3] Departamento Ingeniería Agrícola y Forestal, Universidad de Valladolid, Soria, 42004, Spain
[4] Departamento de Fisica de la Materia Condensada, Universidad de Valladolid, Valladolid, 47011, Spain
[5] Departamento Informatica, Universidad de Valladolid, Valladolid, 47011, Spain
[6] Departamento de Fisica, Universidad de Valladolid, Valladolid, 47011, Spain
[7] Instituto de Ciencias Básica Tecnología e Ingeniería, Universidad Nacional Abierta y a Distancia, Colombia
[8] Departamento de Informática y Sistemas, Universidad de Murcia, 30100, Spain
[9] Departamento Informatica, Universidad de Valladolid, Valladolid, 47011, Spain (Until Febraury 2026)

* **Correspondence:** Emails: thehfmr2011@gmail.com.

Abstract: The estimation of series resistance in photovoltaic (PV) cells is a crucial parameter that significantly influences their efficiency and overall performance. This study proposes a novel methodology to predict the slope of the current–voltage (I–V) curve of a PV cell in the first quadrant, where this slope (the electrical conductance) is directly associated with the series resistance of the cell. By leveraging artificial intelligence techniques, a convolutional neural network model has been developed to estimate this slope from electroluminescence (EL) images of the cells. The model was trained on a dataset consisting of EL images of PV cells with artificially induced defects, together with the corresponding slope values derived from the cells' I–V curves. Furthermore, this work presents a second model that combines the slope information and EL images to improve the prediction of the maximum power point (MPP) of a PV cell, surpassing previous approaches that rely solely on EL imagery. Both models demonstrated low error rates across multiple evaluation metrics, evidencing their accuracy and robustness. Additionally, comparative analysis with other machine learning methods highlights the

competitive performance of the proposed approaches. These contributions provide promising tools for enhancing the assessment and diagnosis of PV cell efficiency and reliability, potentially leading to improved performance and increased longevity of photovoltaic systems.

Keywords: Photovoltaic; electroluminescence; Computer Vision; IV-Curve

1. Introduction

Photovoltaic (PV) cells, the cornerstone of solar power systems, are semiconductor devices that directly convert sunlight into electricity. As smart cities increasingly prioritize sustainable energy solutions, the accurate assessment and modeling of PV cell performance have become one of the most relevant topics in the field. One of the most important techniques for monitoring PV cells is the current–voltage (I–V) curve, which encapsulates the electrical behavior of PV cells under varying operating conditions [1]. This curve provides essential indicators for the efficiency and reliability of solar energy systems, such as open-circuit voltage, short-circuit current, and maximum power point (MPP) [2].

Conventional approaches to I–V curve prediction are grounded in physical modeling, typically based on the single-diode or two-diode models that mathematically represent the behavior of PV cells according to their material and structural properties [3]. However, even though these approaches are well established, they often require detailed information about the manufacturing process and material characteristics that are not usually easily available. Furthermore, their accuracy is frequently constrained by simplifying assumptions and the inherent difficulty of parameter estimation [4].

As the paradigm of smart cities advances, incorporating artificial intelligence, big data analytics, and urban informatics, there is a growing interest in data-driven methodologies to enhance the prediction and monitoring of PV systems. These innovations promise more adaptive and scalable solutions for optimizing energy infrastructure in the context of intelligent urban environments. Machine learning and deep learning models are particularly effective at modeling nonlinear relationships and can generalize from large datasets [5]. By training these models on historical PV data, researchers can estimate I–V curves using other measurable parameters, including environmental and cell-specific inputs, without the need for detailed physical modeling.

Among the parameters derived from the I–V curve, the series resistance (R_s) is a particularly important parameter, as it has a significant impact on the performance and efficiency of PV cells. Series resistance is influenced by the resistance of cell materials and external sources such as contact and interconnection resistances. Estimation of R_s is typically based on analyzing the I–V curve near the open-circuit voltage or through more advanced methods such as impedance spectroscopy [6, 7]. High levels of R_s can result in substantial power losses, lower fill factor, and reduced overall energy conversion efficiency [8].

AI-based approaches to series resistance estimation leverage data-driven models to enhance accuracy and efficiency. While traditional methods rely on analytical techniques and curve fitting, AI models use large datasets and advanced learning algorithms to infer R_s with greater precision [9, 10]. Neural networks trained on diverse I–V curves under various conditions can uncover complex dependencies between input features and R_s values [11]. Deep learning models further enhance performance

by capturing subtle nonlinear patterns that classical methods may overlook [12].

The primary advantage of AI methodologies can be found in their scalability and adaptability. These models accommodate different PV technologies and environmental variations, providing robust and consistent estimations. Furthermore, their performance can improve continuously as additional data becomes available.

An AI-based estimation pipeline typically consists of the following steps: (1) acquisition of a comprehensive I–V dataset under varying operational conditions; (2) preprocessing to remove noise and normalize data; (3) training of machine learning models to capture correlations indicative of series resistance; and (4) inference of R_s for new PV cells based on their I–V profiles [13].

Despite progress in this field, few studies have addressed the prediction of I–V curve parameters directly from visual data, such as images of PV cells. In particular, the estimation of series resistance by analyzing electroluminescence (EL) images has received limited attention. There has been research focused on defect detection using EL images of PV cells and modules [14–16]; other works have used thermographic images [15] or photoluminescence images [17]. More information can be found in reviews on the topic [18, 19]. However, anomaly detection on PV cells is usually based only on the categorization of surface-level issues, without directly considering the effect of those problems on energy production. The focus on predicting the parameters of the I–V curve provides a more direct estimation of the state of the cell. Nonetheless, these methods are not exclusive, as they can be complemented to provide more comprehensive information about the conditions of the modules and cells.

The authors' previous work on the application of AI in the PV field has mainly focused on predicting the performance of PV cells [20, 21], detecting defects in PV modules and cells [22], and generating synthetic images of PV cells [23]. While these works are clearly related to the current study, there are evident differences in focus and methodology.

The main contribution of this paper is the development of a convolutional neural network (CNN)-based model capable of predicting the slope of the I–V curve (the electrical conductance) of a photovoltaic cell from its EL image. This slope is strongly associated with the series resistance, enabling its estimation as a key application. The slope exhibits a more favorable distribution for learning and prediction than direct series resistance estimation, as will be demonstrated in the subsequent sections.

This article constitutes an extended and substantially enhanced version of our previous conference work. Compared to that work, this version provides a more detailed and rigorous description of the methodology, includes additional and recreated figures for clarity, presents an extended experimental evaluation and benchmarking, and, most importantly, introduces a new CNN-based PV performance predictor that exploits the estimated slope as an additional input to improve maximum power point (MPP) estimation, thus demonstrating a practical and previously unexplored application of the predicted electrical parameter.

Another important contribution of this paper, and the main addition in this extended version, is the proposal of a new CNN-based PV performance predictor. This model improves the estimation of the maximum power point (MPP) of the I–V curve by including as input not only the EL image of the PV cell but also the previously computed slope value. This model outperforms existing methods, which are only based on the EL image. The inclusion of the slope value is extremely important since it provides the model with more information for handling cases where the patterns provided only by the EL image are insufficient for accurate MPP estimation. This also highlights the importance of

extracting parameters of the I–V curve, as they provide a detailed view of the performance of PV cells and modules.

This article is an extension of the work presented in ICSC-CITIES 2024 [24]. The data used in this work is open-access, available on Zenodo to download [25].

The remainder of this paper is structured as follows: Section 2 details the proposed methodology. Section 3 presents the experimental results and analysis. Finally, Section 4 concludes the paper and outlines directions for future work.

2. Methodology

This section presents the methodology followed to create models capable of estimating the slope of the I–V curve using the EL image of a PV cell. The process begins with data acquisition and preprocessing. Subsequently, it describes the steps for estimating the I–V curve slope from the training data. Finally, the model architecture and hyperparameter optimization process are presented.

The dataset was divided into three subsets: 60% for training, 24% for validation, and 16% for testing. Only the training and validation sets were used during the training process to avoid data leakage.

2.1. Data acquisition

The data acquisition process involved two independent steps: capturing the EL images of the PV cells using an InGaAs camera, and recording the I–V curves using a cell-level I–V tracer [26]. Further details regarding both acquisition procedures can be found in [21, 23].

The final dataset comprises 852 samples. It includes both monocrystalline and polycrystalline PV cells, and artificial shadows were introduced in several samples to increase variability and improve generalization.

The diversity of the dataset is particularly important, as the inclusion of images with imperfections significantly enhances the model's capacity to generalize to real-world operating conditions. By exposing the model to a wide spectrum of visual patterns—including defects such as degradation, microcracks, surface damage, or soiling—it becomes better equipped to identify and interpret variations in cell behavior under different scenarios. Consequently, the corresponding I–V curve slope values reflect this heterogeneity, exhibiting variability associated with the presence and severity of these defects. This variability is not detrimental; rather, it provides the model with richer information to learn from, ultimately improving its robustness and diagnostic reliability when applied to photovoltaic modules in practical environments.

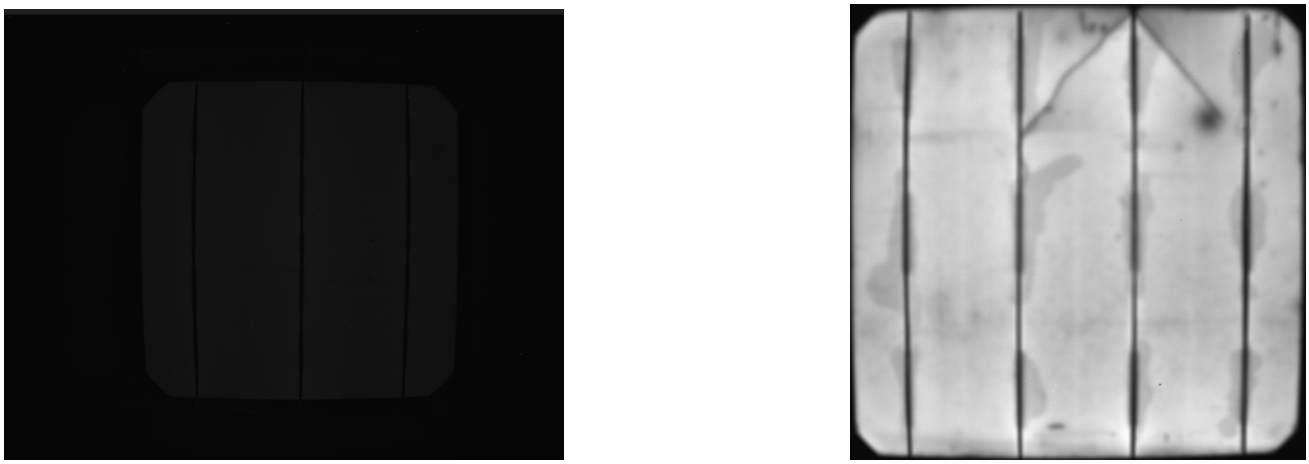
2.2. Image preprocessing

Before training, the EL images were subjected to a series of preprocessing steps, as described in our previous works [21, 23]. These steps include:

- Removal of dead pixels caused by the imaging lens.
- Elimination of luminous noise stemming from image capture artifacts.
- Adjustment of brightness and intensity scale.
- Cropping of surrounding black borders.

- Perspective correction to ensure uniform image geometry.

These operations were implemented in Python. Figure 1 illustrates an example of an EL image before and after preprocessing. More technical details can be found in [23].



(a) Original EL image. **(b)** Preprocessed EL image.

Figure 1. Example of an EL image (a) before and (b) after preprocessing.

2.3. Feature extraction

Traditional machine learning models are not capable of directly handling image data. For this reason, it is necessary to extract features from the images that can adequately describe them. Due to the particular characteristics of the preprocessed EL images, it was necessary to manually determine the features that best represented them. With the help of domain experts, the following features were selected (see also Table 1):

- Statistical metrics such as average, median, mode, variance, and standard deviation values. These features provide a general view of the EL image.
- Proportion of pixels with white values, black values, and other values. These values show the distribution of the intensity of the images. Images with a high number of black values correspond to defective cells.
- Histogram characteristics such as roughness, the number of peaks, the distance between the first two peaks, their height, and their width. These values represent the most important details of the histogram.

Table 1. Selected handcrafted image descriptors utilized for traditional machine learning approaches. Adapted from [21].

Average Intensity	Median Value	Mode Value	Variance
Standard Deviation	Surface Roughness	Dark Pixel Count	Saturated Whites
Peak Count	Peak Distance	Peak 0 Height	Peak 0 Width
Peak 1 Height	Peak 1 Width	–	–

These features have also been used by the research group in other related works [27], where they have demonstrated their ability to represent the images.

With the same objective as the manual extraction system explained before, a feature extraction system based on a pretrained ResNet50 [28] has been tested, using the output before the fully connected layer as the features extracted by the model. This method has been used in the results as an alternative way for testing the feature-based models.

2.4. Slope calculation of I–V curve

The calculation of the series resistance of a PV cell from its I–V curve involves multiple steps. The series resistance can be estimated by analyzing the slope in the final segment of the I–V curve, corresponding to the nearly vertical section near the open-circuit voltage, where the current approaches zero. Figure 2 presents an example of an I–V curve with the three most important points.

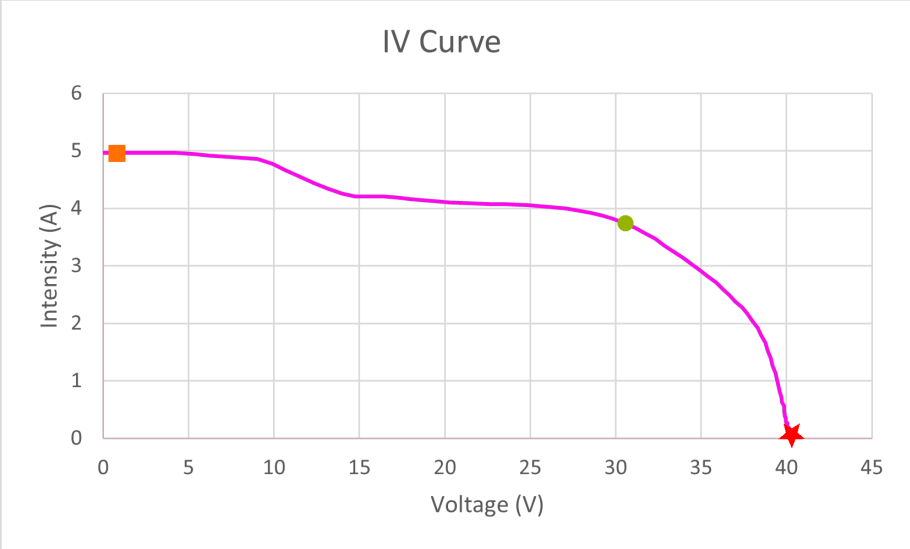


Figure 2. An example of an I–V curve. Red star: open-circuit voltage, orange square: short-circuit current, green circle: maximum power point.

The estimation of the slope is based on a linear regression performed iteratively on the points of the I–V curve, beginning at the open-circuit voltage and incrementally including additional points in reverse order. The slope is calculated in each iteration. In the early steps, slope values exhibit high variability due to the limited number of points and sensitivity to measurement noise.

As more points are incorporated, the slope evolution tends to become smoother. A minimum in the slope evolution (corresponding to the highest negative value) indicates the optimal asymptotic slope. At this point, the effect of noise is minimized, and beyond this point, adding more points decreases the absolute value of the slope (it becomes less negative) due to nonlinearity, leading to greater error in the series resistance calculation.

Figure 3 shows the evolution of the slope as a function of the number of points considered in the regression.

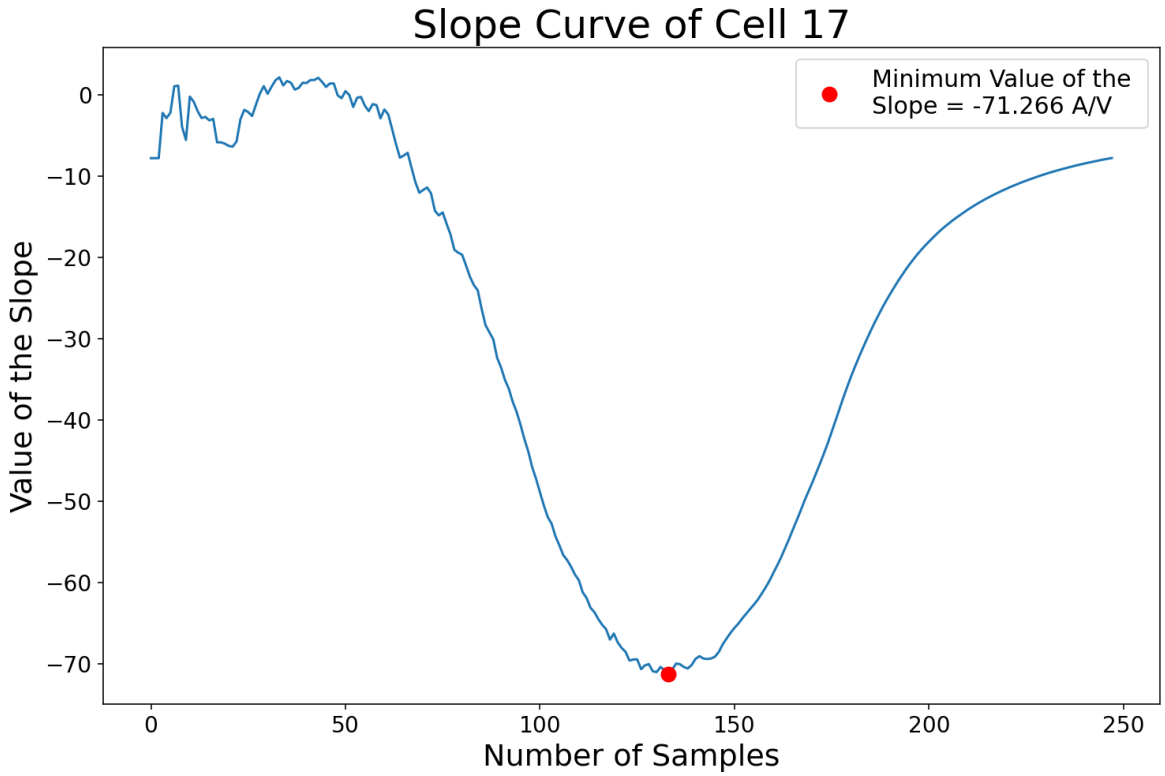


Figure 3. Slope evolution curve according to the number of I–V points used in the linear regression.

The calculated slope quantifies the relationship between voltage and current, corresponding to the electrical conductance of the cell, expressed in siemens (S) according to the International System of Units. This value is reciprocal to the resistance (Ω) [29] , which can be calculated as the inverse of the Slope value.

Figure 4 compares the distributions of the estimated electrical conductance and the calculated series resistance. As shown, the slope distribution is significantly smoother and more regular. On the other hand, the series resistance values display an irregular distribution, making regression more difficult for deep learning models.

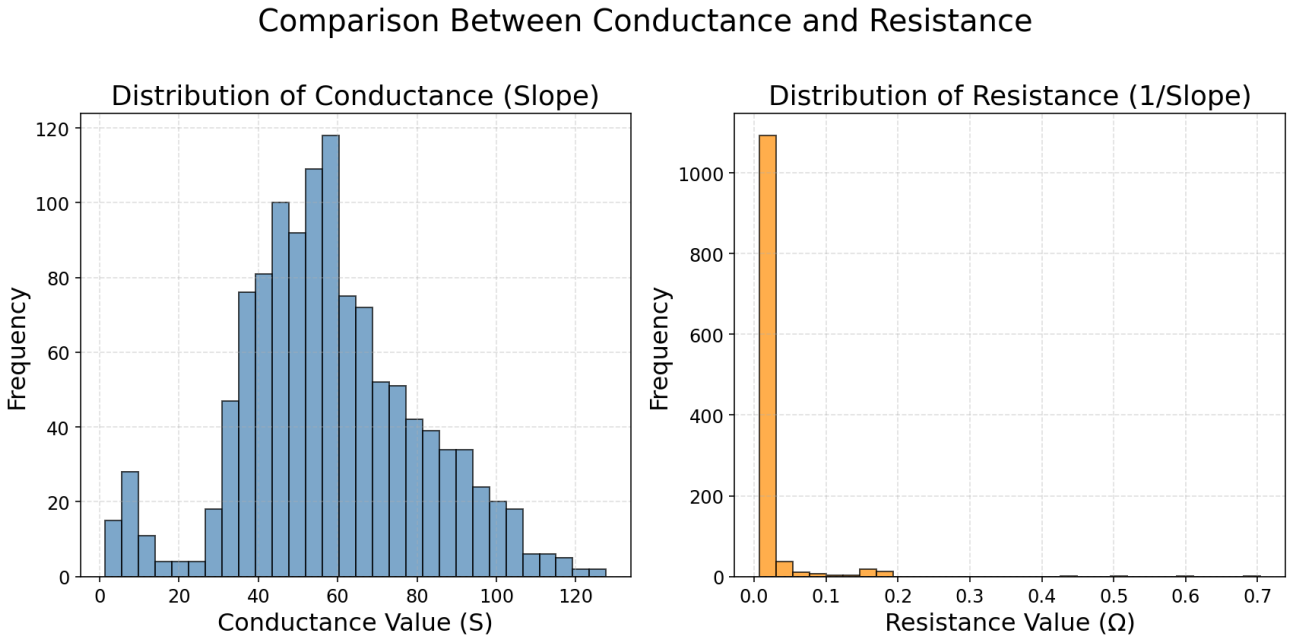


Figure 4. Assessment of the distribution characteristics of conductance (Slopes) versus Series Resistance (1/Slope).

Thus, selecting the electrical conductance (the slope) as the target output provides a better-behaved regression target while still preserving the ability to estimate the series resistance, which can be directly obtained as the reciprocal of the conductance. Nevertheless, it is important to note that the Slope values cannot be standardized, as such transformations would invalidate the resistance calculation.

2.5. Metrics

This section explains the metrics used during the training process. These metrics were applied both to the calculation of the slope and to the estimation of the maximum power point. They were chosen as they are the standard metrics for regression problems.

- Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\text{Predicted}_i - \text{Real}_i| \tag{2.1}$$

- Mean Absolute Percentage Error (MAPE):

$$\text{MAPE}\ (\%) = \frac{1}{n}\sum_{i=1}^{n}\frac{|\text{Predicted}_i - \text{Real}_i|}{\text{Real}_i} \times 100 \tag{2.2}$$

- Mean Squared Error (MSE):

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\text{Predicted}_i - \text{Real}_i)^2 \tag{2.3}$$

- Root Mean Squared Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{Predicted}_i - \text{Real}_i)^2} \tag{2.4}$$

- Root Mean Squared Percentage Error (RMSPE):

$$\text{RMSPE}\ (\%) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\text{Predicted}_i - \text{Real}_i}{\text{Real}_i}\right)^2} \times 100 \tag{2.5}$$

The training and validation processes were performed with these standard metrics, with MAE considered the highest priority metric. The normalized (%) metrics were used for analyzing the models' performance in the testing phase, since they provide a more reliable way of understanding the results. Normalized MSE was not considered, as the squared scale complicates interpretation; instead, RMSE provides equivalent information but in the error domain.

2.6. CNN-based Model

2.6.1. Architecture

The architecture of the CNN was designed through an iterative manual tuning, adjusting the number of layers and parameters based on the considered training metrics. The final model is composed of 18 layers, including combinations of 2D convolutional layers, max-pooling layers, a flattening layer, and fully connected dense layers. The complete architecture is exposed in Figure 5.

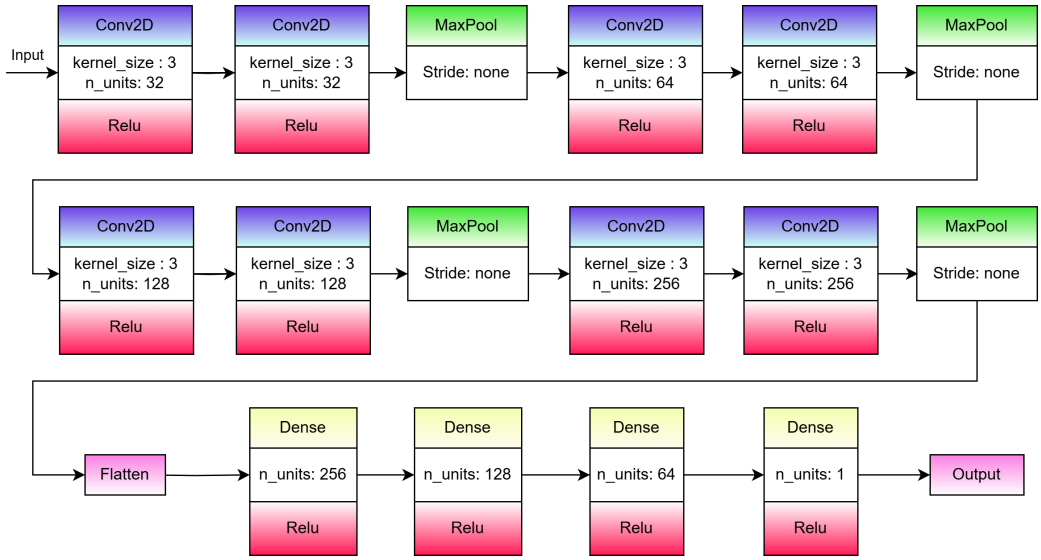


Figure 5. Architecture of the CNN. Adapted from [24].

2.6.2. Hyperparameter Tuning

The performance of CNN is highly dependent on the configuration of its hyperparameters. Given the extensive search space of possible configurations, manual optimization is not feasible. To overcome this limitation, the `Keras Tuner` library [30] has been employed. Specifically, the Bayesian

Optimization strategy [31] was chosen, since it utilizes knowledge from previous evaluations to guide the search for optimal configurations efficiently.

The optimal hyperparameters found through this process of optimization are shown in Table 2.

Table 2. Chosen optimal hyperparameters for the CNN model after tuning within the search space: batch_size $\in \{4, 8, 16, 32, 64\}$, learning_rate $\in [0.01, 0.00001]$, scale_factor $\in [0.5, 1]$, and optimizer $\in \{\text{Adam}, \text{Nadam}, \text{SGD}\}$.

Batch Size	Learning Rate	Scale Factor	Optimizer
4	0.00001	0.5	Adam

2.6.3. Data Augmentation

Given the limited availability of training data, data augmentation techniques were employed to increase the diversity of input samples. Data augmentation [32] enhances the dataset by applying controlled transformations to the original images. In this context, only transformations that preserve the structural integrity of the solar cells were considered, as modifications to their geometry or shape would produce unrealistic instances. The following techniques were applied:

- Horizontal and vertical flips
- Minor rotations

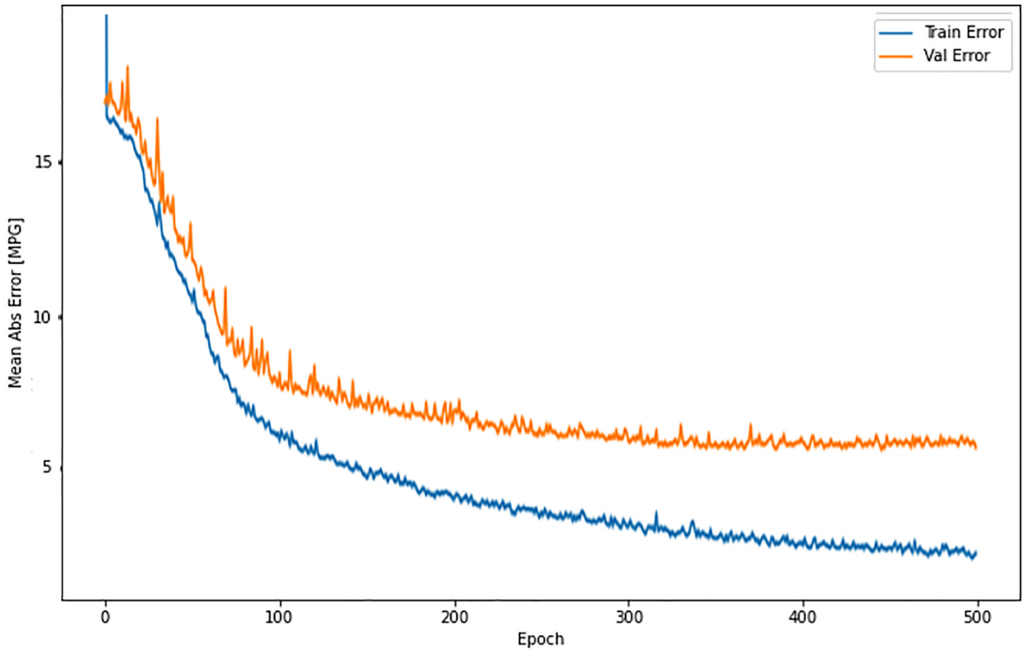
These transformations were applied on-the-fly during training through dedicated preprocessing layers integrated into the network architecture.

2.6.4. Training

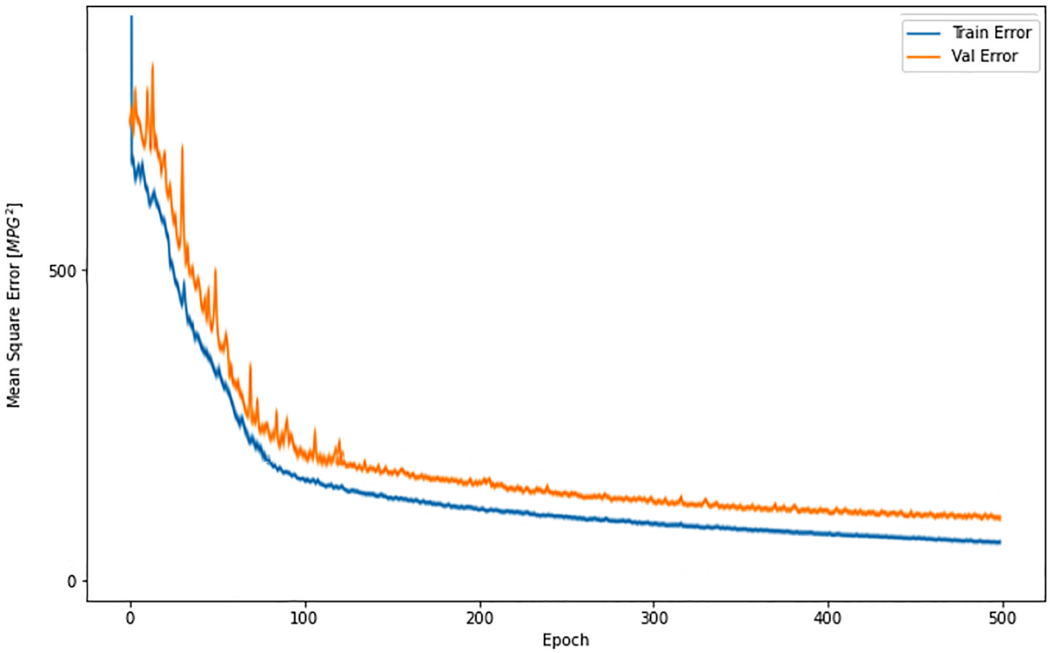
The model was trained using the augmented dataset and the architecture and hyperparameters described in previous sections.

Training was performed during 500 epochs. The metrics considered were Mean Absolute Error (MAE) and Mean Squared Error (MSE). model achieving the lowest on the validation set on both metrics was selected as the final model. Figure 6 presents the evolution of the MAE for both training and validation sets for the best found model. The discrepancies that can be observed between these curves are addressed in the results section.

The model achieved an MAE of 2.110 and an MSE of 25.474 on the training set. On the validation set, the results were an MAE of 5.406 and an MSE of 76.16. Although the optimization process aimed to mitigate overfitting, it is evident that a certain degree of overfitting is still present. However, the model still presents low metric values in Validation and Test set, as it will be presented in section 3. This shows that the model is capable of solving the problem with high reliability even with data not included in the training set.



(a) MAE



(b) MSE

Figure 6. Variation of training and validation metrics over the course of CNN model training.

3. Results

This section evaluates the capabilities of the developed model by analyzing its performance and identifying its potential limitations. Initially, it presents the model performance over the three sets that were used. After that, a comparative analysis is conducted to test the proposed model against other machine learning algorithms. Finally, the estimated slopes are used to improve the estimation of the maximum power point (MPP) of PV cells. The metrics that have been used were explained in Section 2.5.

3.1. Results on different data subsets for the proposed models

As previously described, the dataset is divided into three subsets: Training (60%), Validation (24%), and Testing (16%). Only the Training and Validation sets were utilized during the model development phase, while the Testing set remained completely disjoint to provide an unbiased evaluation of model performance. Table 3 summarizes the results of the proposed model across these subsets.

Table 3. Evaluation results of the developed model on various dataset partitions, measuring MAE, MSE, RMSE, MAPE (%), and RMSPE (%). Adapted from [24].

	Model Performance		
Metric	Training Set	Validation Set	Test Set
MSE	25.47	76.16	62.09
MAE	2.110	5.406	5.041
RMSE	5.047	8.726	7.879
MAPE (%)	1.76	4.51	4.20
RMSPE (%)	4.21	7.27	6.57

The model obtains an MAE of 5.041 on the testing set. Considering the domain of the target variable, this corresponds to a MAPE of 4.2%, which is acceptable within this context. The RMSE and RMSPE metrics also present low values despite being slightly higher than the absolute error case. The reliability provided by the presented metrics shows the possibility of performing accurate detection of PV cells with abnormally high resistance values using only the EL image, without the need for manual inspection.

The results presented in Figure 7 illustrate the variation of the absolute error with respect to the true slope values. Most predictions exhibit errors below 2%. Notably, the largest errors primarily occur for slope values between 20 and 40, which are under-represented in the dataset, as illustrated in Figure 4. This imbalance has a significant impact on the evaluation metrics, indicating that improving the representation of these scarce classes should be a key priority for future research.

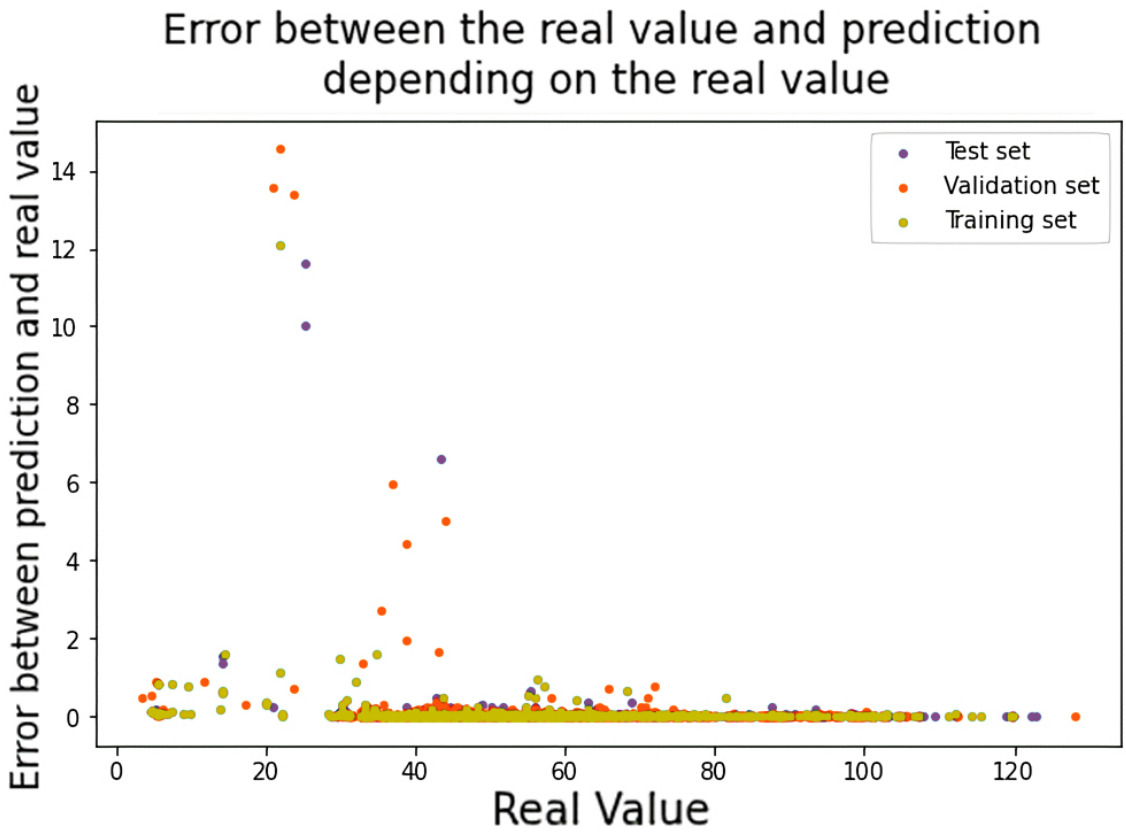


Figure 7. Absolute error of the predicted slope values plotted against their true values. Adapted from [24].

3.2. Comparison with other Machine Learning algorithms

Despite the low error rates presented by the model, it is necessary to perform a benchmarking with other methods. Usually, direct comparisons with previous works in the literature are performed; however, this is not possible in this case due to the lack of studies tackling the issue of directly predicting the I–V curve slope from EL images.

For this reason, this section presents a comparison with various machine learning methods trained using the extracted characteristics described in Section 2.3. These models were designed and trained using MATLAB's Regression Learner App. After performing the training and testing of several models, the four most representative and best-performing models were chosen to be shown in Table 4.

The optimization of the hyperparameters of every model was performed using a Bayesian optimizer during 100 iterations. The final optimized hyperparameters can be found in Table 5. To further improve the comparison, a Deep Neural Network, optimized using the same methodology, has also been included in both Table 4 and Table 5.

Table 4. Comparison of test set performance metrics (MAE, MSE, and RMSE) between the proposed CNN and several machine learning techniques with both feature extraction methods. CNN: Convolutional Neural Network, GPR: Gaussian Process Regressor, SVM: Support Vector Machine, Bagged Tree, NN: Neural Network. The models that use the ResNet50 extractor are indicated with an additional r in their names.

Metric	CNN	GPR	GPRr	LR	LRr	SVM	SVMr	BT	BTr	NN	NNr
MAE	5.041	10.90	7.81	13.07	14.065	12.08	13.18	11.43	11.37	12.80	8.240
MSE	62.09	217.8	129.7	298.5	345.6	280.1	296.6	232.4	240.15	300.65	139.05
RMSE	7.879	14.76	11.38	17.27	19.65	16.73	17.21	15.24	15.49	17.33	11.792
MAPE (%)	0.050	0.109	0.078	0.130	0.140	0.120	0.131	0.128	0.113	0.128	0.082
RMSPE (%)	0.078	0.147	0.113	0.172	0.196	0.167	0.172	0.173	0.153	0.173	0.117

Table 5. Optimal hyperparameters selected for the traditional machine learning models and the neural network.

Algorithm	Main Hyperparameters	Optimal Value(s)
Gaussian Process Regressor	Basis Function	Constant
	Kernel Function	Rational Quadratic
	Isotropic Kernel	Yes
Linear Regression	Terms	Interactions
	Robust Option	off
Support Vector Machines	Kernel	Gaussian
	Kernel Scale	4.4
Bagged Trees	Minimum Leaf Size	8
	Number of learners	30
Neural Network	Hidden layers	[256,128, 64]
	Activation function	ReLU
	Optimizer	Adam (lr = 0.0001)
	Batch size	64

The results indicate that the CNN model substantially outperforms traditional ML algorithms, achieving an MAE of 5.041 compared to 7.81 for the best-performing alternative (GPRr). It is seen how depending on the model, one way of extracting features works better than the other. However, the best performing feature-based method is found with the ResNet50 Features, showing the power of these pretrained methods.

It can be also observed how the neural network model is outperformed by some of the ML traditional methods. Deep Learning methods require high amounts of data and the size of the dataset is not enough to train a feature-based Neural Network. In the case of the CNN, it is enough since it uses directly the image, increasing the amount of available patterns. However, it is worth noting that CNN training demands significantly higher computational resources than other methods.

3.3. *Improvements in estimating the performance of PV Cells*

As explained earlier, the estimation of MPP in the PV field is one of the most interesting problems since it provides a direct measurement of PV cell performance. However, our previous works [27] revealed that MPP estimation is quite vulnerable to unusual values of series resistance. The tested methods were not capable of correctly predicting the MPP in PV cells with extremely high series resistance values.

For this reason, this section presents an improved MPP estimator, based on the approach presented in our previous works [27]. This new model is used to evaluate the practical benefit of estimating the I–V curve slope as an indicator of the series-resistance value. The I–V slope is used as an additional input to the model.

The improved model retains a CNN-based architecture, depicted in Figure 8. The new input is included after the Flatten layer, since it is a scalar value, not a 2D image. Training of this network was performed with the hyperparameters listed in Table 6. These parameters were chosen following the same methodology explained in Section 2.6. Training, validation, and testing were performed with the same dataset used in the original article.

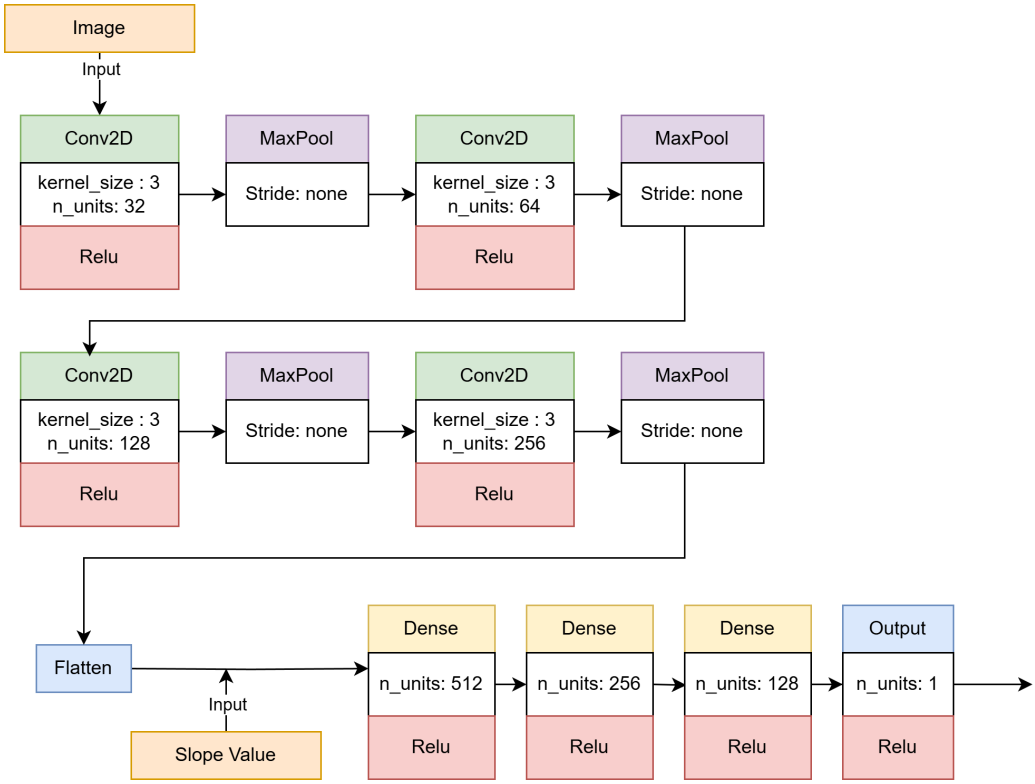


Figure 8. Architecture of the Convolutional Neural Network (CNN).

Table 6. Optimal hyperparameters found for the Convolutional Neural Network: batch_size $\in [8, 16, 32, 64]$, learning_rate $\in [0.01, 0.00001]$, scale_factor $\in [0.5, 1]$, optimizer $\in$ {Adam, Nadam, SGD}.

batch_size	learning_rate	scale_factor	optimizer
8	0.00001	0.66	Adam

Training The chosen metrics were the Mean Absolute Error (MAE) and the Mean Squared Error (MSE).

Table 7 compares error metrics among several models: the enhanced model (*Resmodel*), the original model from the cited paper (*Origmodel*), and additional models that were tested. These feature-based models were trained with the features explained in Section 2.3.

Table 7. Performance comparison on the test set between the proposed improved estimator, the original estimator, and various machine learning algorithms: Linear Regression (LR), Support Vector Machines (SVM), and Gradient Boosting Regressor (GBR). Metrics reported are MAE and MSE.

Metric	Resmodel	Origmodel	RF	SVM	GBR
MAE	0.0231	0.0262	0.0712	0.0686	0.0681
MSE	0.0009	0.0011	0.0095	0.0103	0.0092

The results provide evidence that including the estimated slope as an input improves the precision of the estimation, since it provides relevant additional information that cannot be extracted directly from the image. The improved method outperforms the original model in both metrics considered in the original paper (MAE and MSE).

These results also demonstrate the practical utility of the slope of the I–V curve as an indicator of the series resistance of a PV cell. Future work will focus on other issues that can benefit from incorporating this value.

4. Conclusions and future work

The monitoring of PV modules and cells is a critical research area in the PV field, as it provides valuable information about their productivity. The series resistance affecting the cells significantly impacts their performance. This effect can be easily observed in the I–V curves, where the slope depends on the series resistance. To estimate the series resistance without directly measuring the I–V curve, a model capable of analyzing the EL images of PV cells has been developed. This model was trained with pairs of I–V curves and EL images. It can predict the slope of the I–V curve, which can then be used to calculate the series resistance.

The results indicate that the model predicts the slope with an MAE of approximately 5 and an RMSE of 7.8, outperforming other machine learning methods such as Gaussian Process Regressor, Support Vector Machines or Neural Networks

However, some limitations remain. Most prediction errors occur in cells with slope values between 20 and 40, which are under-represented in the dataset. This could be fixed by two different approaches:

By measuring new PV cells with I–V Slope values focused on those intervals or by creating synthetic data using Generative AI techniques.

The training of CNN models is computationally intensive due to their complexity and memory requirements. Furthermore, real-time application of the current approach is limited by the need to disconnect the PV module, which prevents instantaneous acquisition of EL images. To address this limitation, non-intrusive imaging modalities, such as photoluminescence or thermography captured by drones, will be investigated to enable real-time deployment, offering significant potential for large-scale solar farm monitoring.

Extending this methodology to module-level diagnosis requires a substantial number of PV modules characterized with both EL images and I–V curves. Most open-access datasets provide only EL images and lack the corresponding I–V curves necessary to generate the training labels. A key aspect of our future work will therefore focus on acquiring a large, high-quality dataset of PV modules including both types of measurements.

Future work will also focus on exploring more advanced deep learning architectures, such as ResNet and EfficientNet, not only as feature extractors but also to evaluate their potential for directly estimating key parameters such as the series resistance and the maximum power point. Moreover, the integration of large language models will be investigated to assess their potential in enhancing data interpretation, model generalization, and automated diagnostic capabilities within photovoltaic analysis.

This paper has also shown that including this predicted value as an input to a CNN-based output power estimator improves its predictive performance, lowering the MAE and MSE values.

To sum up, the proposed model offers an effective approach for estimating the slope of PV cells. However, further research is required to enhance the dataset, explore applicability in real-time scenarios, and investigate the performance of other techniques that have been applied in similar fields, such as Transformers or Generative AI.

Funding

This research was supported by the Universidad de Valladolid through the 2020 predoctoral contracts, co-funded by Santander Bank. Additional funding was provided by the Spanish Ministry of Science, Innovation, and Universities within the framework of the "Plan Estatal de Investigación Científica, Técnica y de Innovación" (project ID: PID2023-148369OB-C43). Further financial support was received from the Spanish Ministry of Science and Innovation under project PID2020-113533RB-C33, as well as from the Fundación Séneca, Agencia de Ciencia y Tecnología de la Región de Murcia (project 22130/PI/22). The Universidad de Valladolid also supported this work through the ERASMUS+ KA-107 program. The authors additionally acknowledge the MOVILIDAD DE DOCTORANDOS Y DOCTORANDAS UVA 2024 program at the University of Valladolid for their support.

Use of AI tools declaration

The authors declare that no Artificial Intelligence (AI) tools were used in the writing or analysis of this article.

Conflict of interest

All authors declare no conflicts of interest in this paper. Dr. Luis Hernández-Callejo is an special issue editor for Mathematical Biosciences and Engineering and was not involved inthe editorial review or the decision to publish this article.

References

1. S. M. Sze, K. K. Ng, *Physics of Semiconductor Devices*, Wiley, Hoboken, NJ, 3rd edition, 2007.
2. M. A. Green, *Solar Cells: Operating Principles, Technology, and System Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1982.
3. A. Jain, A. Kapoor, Exact analytical solutions of the parameters of real solar cells using lambert w-function, *Solar Energy Mater. Solar Cells*, **81** (2004), 269–277. https://doi.org/10.1016/j.solmat.2003.11.018
4. D. S. H. Chan, J. C. H. Phang, Analytical methods for the extraction of solar-cell single- and double-diode model parameters from i-v characteristicsm, *IEEE Transact. Electron. Devices*, **34** (1987), 286–293. https://doi.org/10.1109/T-ED.1987.22920
5. Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature*, **521** (2015), 436–444. https://doi.org/10.1038/nature14539
6. M. A Green, *Solar Cells: Operating Principles, Technology, and System Applications*. Prentice-Hall, 1982.
7. M. Wolf, H. S. Rauschenbach, Series resistance effects on solar cell measurements, *Adv. Energy Convers.*, **3** (1963), 455–479. https://doi.org/10.1016/0365-1789(63)90063-8
8. Y. Hishikawa, T. Kinoshita, K. Tanaka, T. Nishiwaki, Series resistance determination from current-voltage characteristics at different illumination intensities, *Progress Photovolt. Res. Appl.*, **16** (2005), 465–475.
9. J. Lee, J. Park, K. Yoon, Machine learning for solar cell design and characterization, *J. Power Sources*, **451** (2020), 227743. https://doi.org/10.1016/j.jpowsour.2020.227743
10. Z. Chen, Y. Liu, Y. Yang, Artificial intelligence-based framework for photovoltaic system fault diagnosis and performance prediction, *Renew. Sustain. Energy Rev.*, **136** (2021), 110422.
11. P. Mehta, R. Ramprasad, Applying machine learning to accelerate material discovery for photovoltaic applications, *J. Appl. Phys.*, **127** (2020), 190901.
12. W.He, L.Zhang, M. Chen, Data-driven methods for fault detection in photovoltaic arrays, *Solar Energy*, **193** (2019), 329–343.
13. K. Patel, S. K. Jha, R.Rajput, Neural network-based approach for PV system performance prediction, *Renew. Energy*, **172** (2021), 93–104.
14. S. Nara, T. Fuyuki, et al. Quantitative prediction of power loss for damaged photovoltaic modules using electroluminescence, *Energies*, **11** (2018), 1172. https://doi.org/10.3390/en11051172
15. T. Sauer, M. Rinio, U. Eitner, J. Bagdahn, Characterization of series resistance in photovoltaic modules by combined electroluminescence and infrared thermography imaging, *Solar Energy Mater Solar Cell*, **95** (2011), 2131–2135.

16. O. Breitenstein, M. Langenkamp, Series resistance imaging of solar cells by voltage dependent electroluminescence, *Progress Photovolt. Res. Appl.*, **11** (2003), 515–526. https://doi.org/10.1002/pip.520

17. T. Fuyuki, H. Kondo, M. Fukawa, A. Kitiyanan, Y. Yamazaki, Quantitative evaluation of electroluminescence images of solar cells, *Japanese J. Appl. Phys.*, **45** (2006), L845.

18. H. F. Mateo Romero, M. A. Gonzalez Rebollo, V. Cardenoso-Payo, V. Alonso Gomez, A. Redondo Plaza, R. T. Moyo, et al., Applications of artificial intelligence to photovoltaic systems: A review, *Appl. Sci.*, **12** (2022). https://doi.org/10.3390/app121910056

19. A. Mellit, S. A. Kalogirou, Artificial intelligence techniques for photovoltaic applications: A review, *Progress Energy Combust. Sci.*, **34** (2008), 574–632. https://doi.org/10.1016/j.pecs.2008.01.001

20. H. F. Mateo-Romero, M. E. Carbonó dela Rosa, L. Hernández-Callejo, M. González-Rebollo, V. Cardeñoso-Payo, V. Alonso-Gómez, et al., Enhancing photovoltaic cell performance evaluation: An adaptive neural fuzzy inference modeling approach, *Progress in Artificial Intelligence*, 2024, under review.

21. H. F. Mateo-Romero, L. Hernandez-Callejo, M. A. G. Rebollo, V. Cardeñoso-Payo, V. A. Gomez, J. I. M. Aragonés, et al., Optimized estimator of the output power of PV cells using el images and i–v curves, *Solar Energy*, **265** (2021), 112089. https://doi.org/10.1016/j.solener.2023.112089

22. H. F. Mateo-Romero, Á. Pérez-Romero, L. Hernández-Callejo, S. Gallardo-Saavedra, V. Alonso-Gómez, J. Morales-Aragonés, et al., Photovoltaic cells defects classification by means of artificial intelligence and electroluminescence images, In: Sergio Nesmachnow and Luis Hernández Callejo, editors, *Smart Cities*, pages 31–41, Cham, 2022. Springer International Publishing. https://doi.org/10.1007/978-3-030-96753-6_3

23. H. F. Mateo-Romero, L. Hernandez-Callejo, M. A. G. Rebollo, V. Cardeñoso-Payo, V. A. Gomez, H. J. Bello, et al., Synthetic dataset of electroluminescence images of photovoltaic cells by deep convolutional generative adversarial networks, *Sustainability*, **15** (2023), 7175. https://doi.org/10.3390/su15097175

24. H. F. Mateo-Romero, J. I. M. Aragonés, L. Hernández-Callejo, M. Á. González-Rebollo, V. Cardeñoso-Payo, V. Alonso-Gómez, et al., Novel convolutional neural network model for estimating series resistance in PV cells by predicting i-v curve slopes on electroluminescence images, In: Sergio Nesmachnow and Luis Hernández Callejo, editors, *Smart Cities*, pages 105–117, Cham, 2025. Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-85324-1_8

25. H. F. Mateo-Romero, J. I. Morales, L. Hernández Callejo, V. Cardeñoso-Payo, M. Gonzalez, V. Alonso Gómez, et al., Dataset of paper "cnn-based estimation of series resistance in photovoltaic cells from electroluminescence images with application to output power prediction", September 2025. https://doi.org/10.5281/zenodo.17083852

26. J. I. Morales-Aragonés, V. A. Gómez, S. Gallardo-Saavedra, A. Redondo-Plaza, D. Fernández-Martínez, L. Hernández-Callejo, Low-cost three-quadrant single solar cell i-v tracer, *Appl. Sci.*, **12** (2022). https://doi.org/10.3390/app12136623

27. H. F. Mateo-Romero, *Employing artificial intelligence techniques for the estimation of energy production in photovoltaic solar cells based on electroluminescence images*, PhD thesis, Universidad de Valladolid, 2024.

28. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016. https://doi.org/10.1109/CVPR.2016.90

29. D. A. Bonnell, R. J. Nemanich, *Characterization of Electronic Materials and Devices*, Wiley-VCH, Weinheim, Germany, 2000.

30. T. O'Malley, E. Bursztein, J. Long, F. Chollet, H. Jin, L. Invernizzi, et al., Keras Tuner, 2019. Available from: `https://github.com/keras-team/keras-tuner`

31. R. Garnett, *Bayesian Optimization*, Cambridge University Press, 2022. https://doi.org/10.1017/9781108348973

32. C. Shorten, T. M. Khoshgoftaar, A survey on image data augmentation for deep learning, *J. Big Data*, **6** (2019), 60. https://doi.org/10.1186/s40537-019-0197-0

AIMS Press