Estimation of the performance of Photovoltaic Cells by means of an Adaptative Neural Fuzzy Inference model

Hector Felipe Mateo-Romero[1][0000−0002−5569−3532], Mario Eduardo Carbonó dela Rosa[2][0000−0002−6741−6145], Luis Hernández-Callejo[1][0000−0002−8822−2948], Miguel Ángel González-Rebollo[1][0000−0001−9130−7339], Valentín Cardeñoso-Payo[1][0000−0003−1460−158X], Victor Alonso-Gómez[1][0000−0001−5107−4892], Óscar Martinez-Sacristán[1][0000−0002−2283−0350], and Sara Gallardo-Saavedra[1][0000−0002−2834−5591]

[1] Universidad de Valladolid, Spain,
H.F.M.R. `hectorfelipe.mateo@uva.es`
L.H.C. `luis.hernandez.callejo@uva.es`
M.A.G.R. `mrebollo@eii.uva.es`
V.C.P. `valen@infor.uva.es`
V.A.G. `victor.alonso.gomez@uva.es`
O.M.S `oscar.martinez@uva.es`
S.G.S. `sara.gallardo@uva.es`
[2] Universidad Nacional Autónoma de México `mecr@ier.unam.mx`

Abstract. This paper presents an Adaptive Neuro-fuzzy Inference System capable of predicting the output power of photovoltaic cells using their electroluminescence image and their IV curve. The input consists of 3 different features: the number of black pixels, grey pixels and white pixels. ANFIS combines the learning capabilities of Artificial Neural Networks with the comprehensible rules of Fuzzy Logic, being optimal for this problem, as demonstrated by the metrics of MAE of 0.064 and MSE of 0.009, which are better than the performance of other tested methods such as Support Vector Machines or Linear Regressor.

Keywords: Fuzzy Logic · Photovoltaic · Electroluminescence · Machine Learning · ANFIS

1 Introduction

Energy is the motor of every sector, being fossil fuels the most important source of energy, with an 80.2% in 2019 [17]. This trend is changing in recent years with the inclusion of renewable energies. Different problems such as wars, and climate change are provoking a shift in favor of these kinds of energies. Of the different kinds of renewable energies, solar energy is one of the most important ones for smart cities since solar panels can be easily installed in buildings.

Photovoltaic (PV) energy is produced by PV panels using the energy from the sun. These modules are composed of a big amount of small units known as PV cells. They can suffer from different conditions [1] that can affect their performance and security. Constant monitoring of their condition is vital when the optimization of their production is needed.

The monitoring and maintenance of PV installations have been traditionally a manual process but this is not feasible when the size of the facilities reaches high dimensions or when they are included at places of difficult access. Artificial Intelligence (AI) takes an important place in this field since AI techniques can be used to improve the production and control the conditions of the modules [11, 13].

The technique known as Electroluminescence (EL) [5] is the most used technique to capture the surface of PV modules and cells, consisting in capturing the light emitted by the PV units when they are being injected with electric current. Different works have used these images to detect the defects in the surface of PV cells [16]. Of the different AI techniques, Convolutional Neural Networks [2] are usually the best performing, but they have some limitations since it is a non-symbolic method, meaning that no knowledge can be extracted from the reasoning of the method.

This paper presents a new approach to analyzing the state of PV cells. The data is not only obtained with the EL techniques but also with the information about the energy produced by the PV cell by measuring their Intensity-Voltage IV curve. Another important feature of this paper is that it applies a combination of two different methods (Fuzzy Logic and Neural Networks) known as Adaptative Neuro-fuzzy Inference System (ANFIS) [8]. This combination is capable not only of obtaining a good performance in the problems, using the capacities of Neural Networks but also to obtain understandable knowledge thanks to the symbolic aspect of the fuzzy logic.

Other PV problems have been tackled in other problems using Fuzzy logic or ANFIS systems for detecting microcracks [4], modeling the PV systems [14] or finding the Max Power Point [9]. ANFIS has been also applied to forecasting problems [6]. However, no work has dealt with our exact issue, of finding the performance of a PV cell based on the IV curve using the EL images.

The rest of the paper is organized as follows: Section 2 explains the basics of ANFIS, Section 3 explains the methodology used, Section 4 shows the results and the conclusions that can be drawn from them, finally Section 5 presents the conclusions of the paper.

2   ANFIS

The Adaptative Neuro Fuzzy Inference System [7] links Artificial Neural Networks (ANN) [3] with Fuzzy Logic (FL). Its inference system is based on the Takagi-Sugeno fuzzy logic [18] with IF-THEN rules. The combination of ANN and FL creates a model that is capable of updating under new situations thanks

to the training function of the ANN. The basic architecture of the network(See Fig. 1) is composed of 5 different layers :

- Fuzzification layer: In this layer, the parameters of the input membership functions are determined, for example: number, range, and the type of each membership function (triangle, trapezoid, generalized bell, and Gaussian), among others. Moreover, each node (information processing points) is adaptive.
- RRules layer: In this layer, the knowledge rules that relate the inputs and outputs established in the system are created. It is composed of fixed nodes that represent the firing power of each rule.
- Normalization Layer: The function of this layer is to normalize the firing strength of each rule by dividing it by the sum of the firing strength of all rules.
- Defuzzification Layer: This layer produces a weighted output of each rule using the normalized values.
- Output Layer: The output of this layer is the sum of all its inputs and provides the overall output of the network.
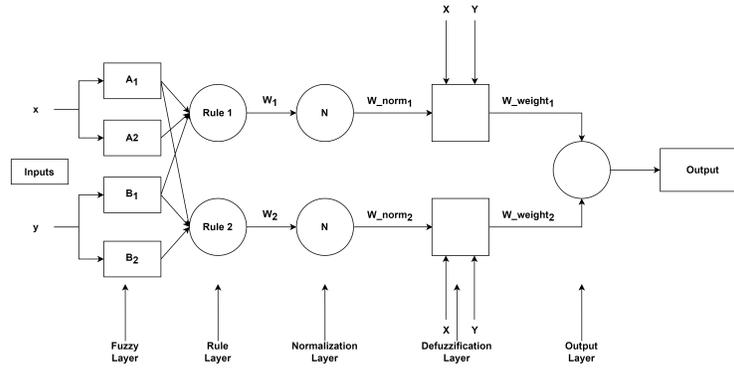

Fig. 1: Structure of a generic ANFIS model.

3 Methodology

This section explains the different procedures that were performed to prepare the model. This includes the collection, labeling, and preprocessing of the data, as well as the creation and optimization of the ANFIS model.

3.1 Data gathering

The retrieval of the data was divided into two different processes, a more detailed explanation can be found [12]:

– Electroluminescence images: These images were captured using an "InGaAs C12741-0" silicon detector camera with an 8 mm focal length lens and an f-number of 1.4 in the same temperature conditions. The images were taken in total solar shielding to reduce light noise. Different irradiance levels were used to obtain more images.
– IV Curve: They were obtained using a 3-quadrants IV-tracer [15], obtaining only the information of the active zone for the experiments. Fig. 2 presents different IV curves for the same cell.
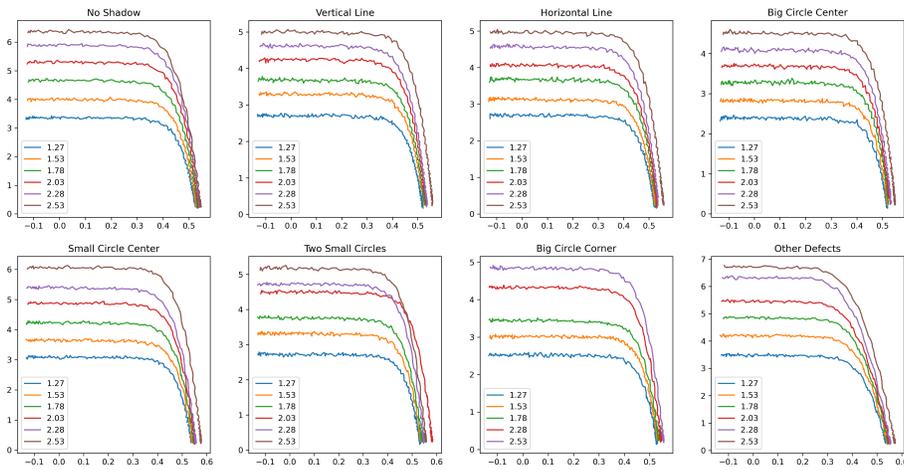

Fig. 2: Different IV curves for the same cell. Each diagram represents a different defect, and each line a IV curve measured a certain irradiance. The label indicates the current used

The obtained dataset is composed of 666 EL images with their corresponding IV curve.

3.2   Image Preprocessing

The images presented some problems related to the noise of the lights, the scale of the lights in the images, the black contours presented in the images, and their perspective. These problems are harmful for visual inspection and also for Artificial Intelligence algorithms since it is more difficult for them to find patterns in the images. For this reason, a preprocessing has been performed on the images removing the luminous noise, standardizing the histogram, and finally cropping the image to the cell surface. (See Fig. 3), more information can be seen in [12].
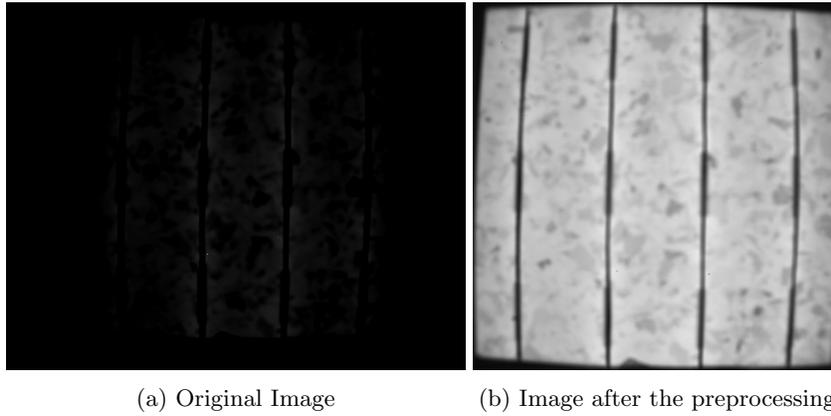
(a) Original Image   (b) Image after the preprocessing

Fig. 3: A sample image before and after preprocessing.

3.3   Maximum Power Normalization

The IV curve provides information about the amount of energy produced by each PV cell. This information cannot be obtained directly: First, the IV curves are separated according to the irradiance used to take the measurement, and then according to the Max Power Point of the curve.
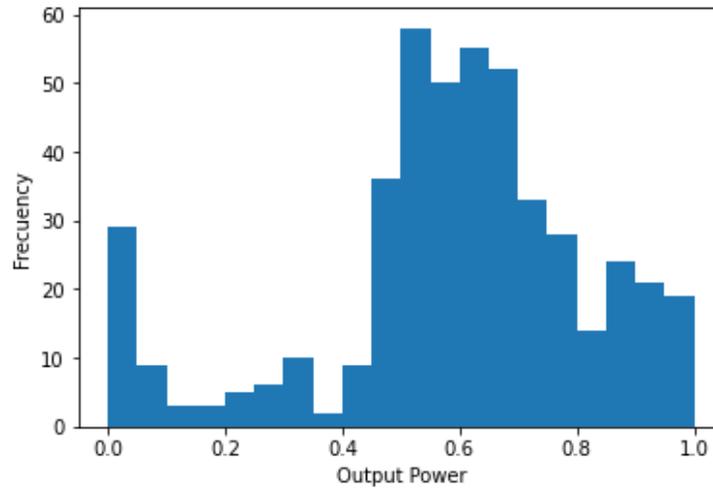

Fig. 4: Architecture of the model

Finally, two different normalizations are performed:A z-score normalization using the mean and standard deviation, and a min-max normalization using

the maximum and minimum values. The objective of this normalization is to obtain an output that is independent of the level of irradiance used for taking the measure.

The variable obtained after the multi-normalization has a range of $[0,1]$, with low values corresponding to cells with poor performance and high values to cells with good performance. The histogram of the distribution of the power values can be seen in Fig. 4, which shows the values over the whole domain. This variable will be considered as the output of the system.

3.4   Feature Extraction

The ANFIS architecture is built using ANNs, which are not suitable for dealing with images directly [10]. The most common solution to this problem is to extract features from the images. These features describe the most important details of the images.

The data set is divided into 3 different sets: Training (70%), Validation (15%), and Testing (15%). Three different features (Blacks, Grays, and Whites), which will be the inputs of the model, are extracted from the images using their color histogram, previously normalized between 0 and 1. These features are obtained by dividing the histogram into different subgroups and obtaining the number of pixels in each group. The boundaries between the different subgroups are computed using the following steps:

- The color histogram of each image is obtained and accumulated in order to general histogram of all images

- The minimum point between the first and second peaks is used to separate the black area from the gray area.

- The minimum point between the second and third peaks is used to separate the gray area from the white area.

- The mean between all of the values of each minimum point is computed. These values are rounded an used for the limit of each group (0.35 and 0.7). Fig 5 presents a cell and its histogram where 3 different peaks can be seen, the divisions between the three groups are also presented.

- The amount of pixels of each group is divided by the total amount of pixels of each image.
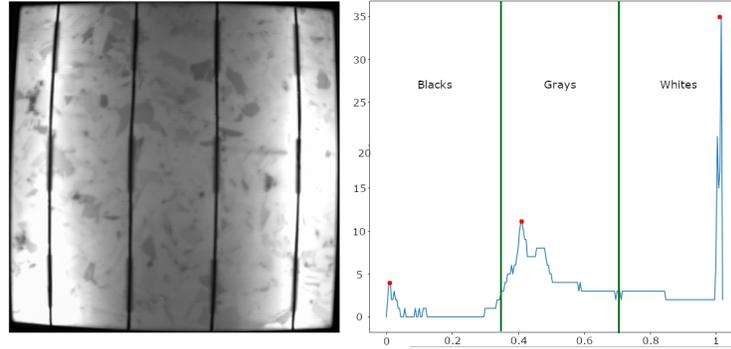
Fig. 5: A PV cell and its histogram. Red dots represent the peaks. The green lines represent the division between the groups for calculating the features

3.5   Model and Architecture

As explained before, the method chosen to solve this problem is ANFIS, since it combines the powerful learning capabilities of ANN with the comprehensible rules of FL. The ANFIS model was implemented with the Neuro-Fuzzy Designer application of Matlab. The diagrams and graphics were also obtained with that application.

The design of the model architecture was chosen based on the best performance of the mean absolute error (MAE), mean square error (MSE) and root mean square error (RMSE) in the training, testing and validation stages. For this it was necessary to run multiple tests varying the training epochs, the number and type of input membership functions. Taking care not to overfitting the model to obtain an acceptable number of rules that would not generate additional computational expense.

The optimization of the membership functions for the inputs has consisted of a combination of a manual process (The selection of the shape of the functions and the number of membership functions for each input) and an automatic process during the training of the network (The size of the intervals and the intersections between them). Different shapes for the membership functions such as Triangular Functions, Trapezoidal Functions, Phi-Shaped, Simple Gaussian Functions, and Double Gaussian Functions) were tested, being the Double Gaussian function the one that produced the best performance in the model. In a similar way, a different number of membership functions were tested for the three inputs, obtaining that two for the black input, four for the gray input, and five for the white input were the optimal values

The final model contains 40 if-then rules due to the combination of all different input membership functions ($2 * 4 * 5 = 40$). Figure 6 shows the architecture of the model.
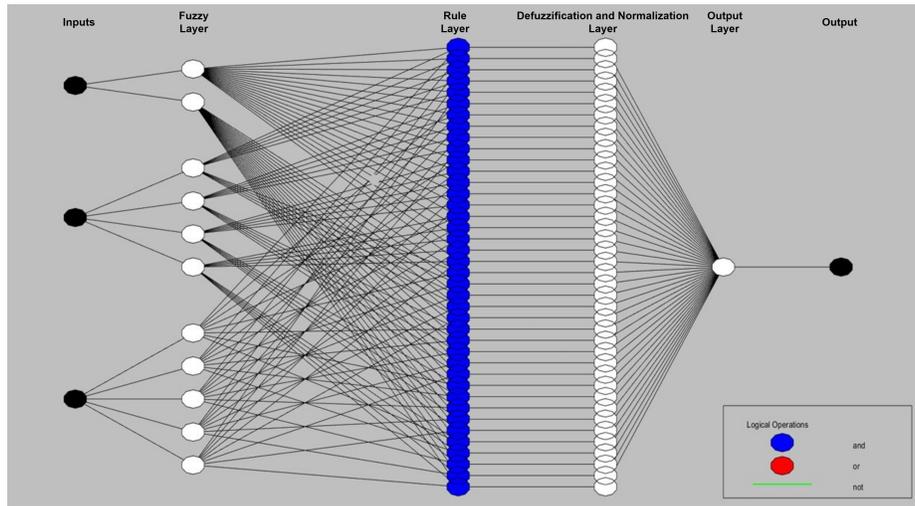
Fig. 6: Architecture of the model

3.6   Training

The developed model must be trained using the data from the training and validation sets. Each sample has 3 inputs and 1 output. As mentioned before, there are 666 samples, divided into 70%, 15%, and 15%. The optimization is done with a combination of backpropagation for the input membership functions and least squares for the outputs, trying to minimize the Mean Squared Error for the output. Fig. 7 shows the evolution of the training and validation errors over 1000 epochs.
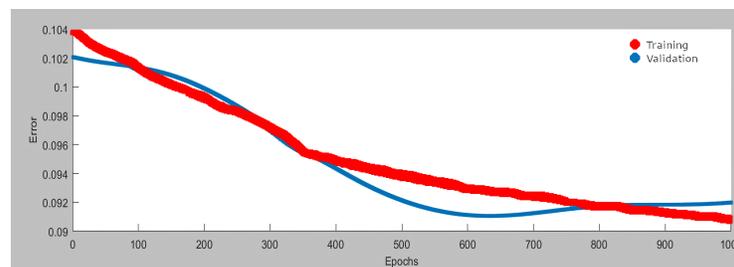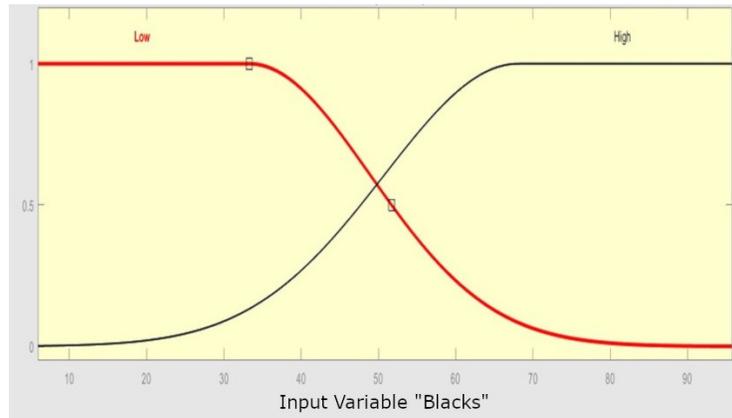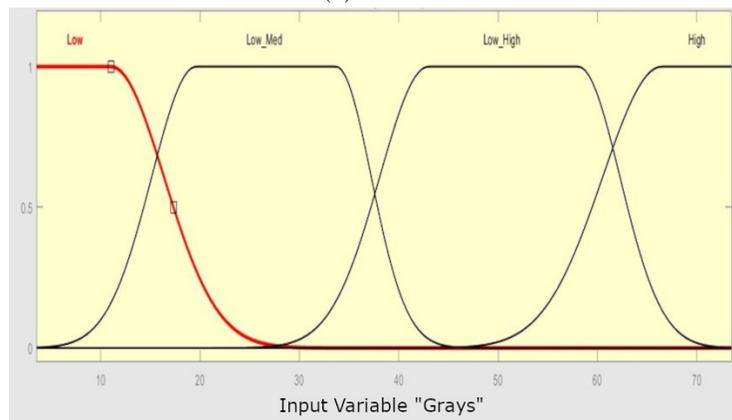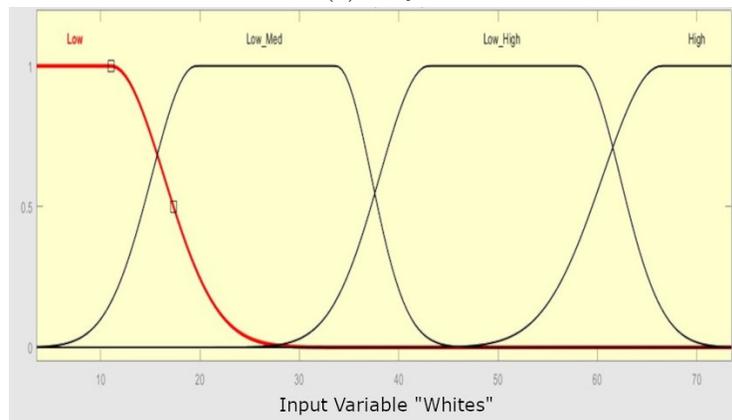

Fig. 7: Evolution of the training error and Validation error during the training process

(a) Blacks

(b) Grays

(c) Whites

Fig. 8: Membership functions of the three different inputs. These functions have been obtained by the combination of manual optimization and the automatic training of the ANFIS method.

It can be seen that around epoch 600 the validation error is minimized (0.0911), since after this epoch the error increases steadily, a clear indication of an overfitting problem when training continues for more than 600 epochs.

Fig. 8 shows the membership functions of the inputs after their optimization in the training process.

3.7   Rules

As explained before, the rules are obtained automatically during the training process, Fig. 9 shows the 40 rules of the system, for every possible combination of inputs.
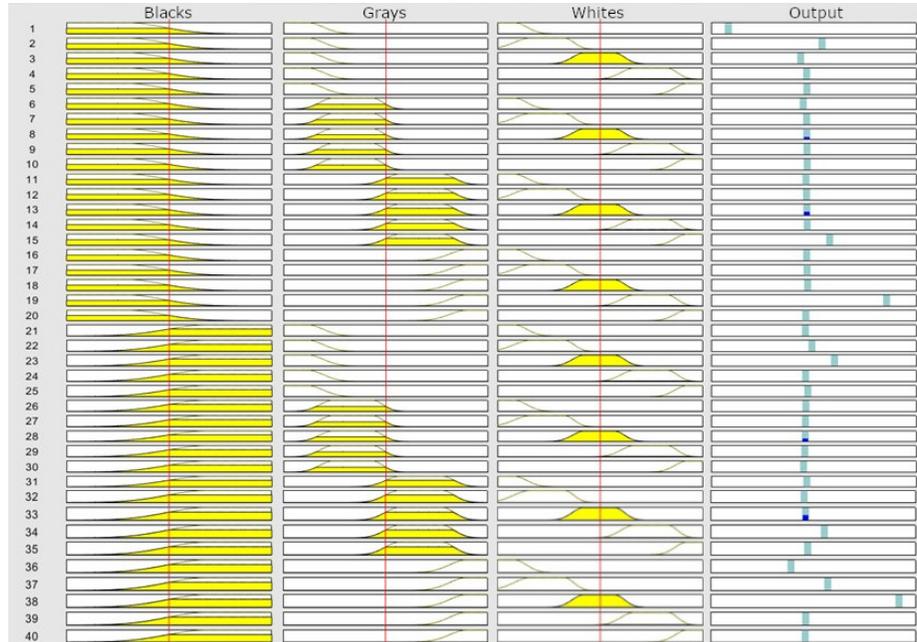

Fig. 9: Fuzzy Rules

In ANFIS and Fuzzy Logic systems, it is typical to present 3D graphics that show the relationship between the variables and as a graphical representation of the rules, showing the effects of the rules in the output. Fig. 10 shows the effects of the variables black and gray in the output. It can be how the different values of both variables change notably the output.
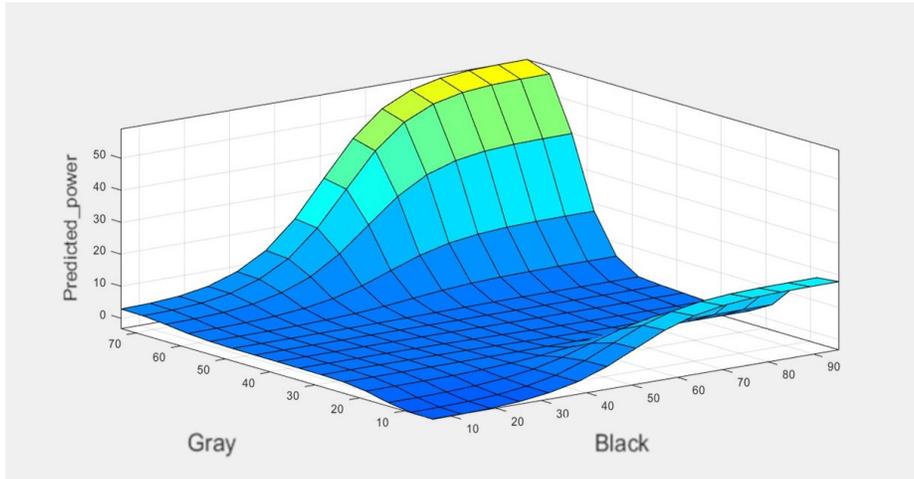
Fig. 10: 3D surface diagram of the effects of Gray and Black on the output

4 Results

This section tests the capabilities of the created model to show its performance and find its weaknesses. First, the performance is checked with different sets. Then the performance of the model is compared with other machine learning models.

4.1 Metrics in the different sets

Table 1 shows the performance of the model with different metrics. it can be observed how the performance of the model is stable in the three sets, with a minor decrement in the testing set, which is expected since that data was completely unrelated to the training process.

Table 1: Results according to the model of the model in the different sets. MAE: Mean Absolute Error, MSE: Mean Squared Error, RSME: Root Square Mean Error

Metric	Training	Validation	Testing
MAE	0.05878	0.06416	0.06474
MSE	0.00871	0.00830	0.00940
RSME	0.09297	0.09112	0.09695

Fig. 11 presents the diagram of the outputs for the training and validation sets. It shows how in most cases the actual value and the predicted value are

extremely close, this is a shred of clear evidence that the model's performance is good. Moreover, the number of outliers that have a high distance from their label is considerably low, and even most of them continue to maintain the essence of the actual value.
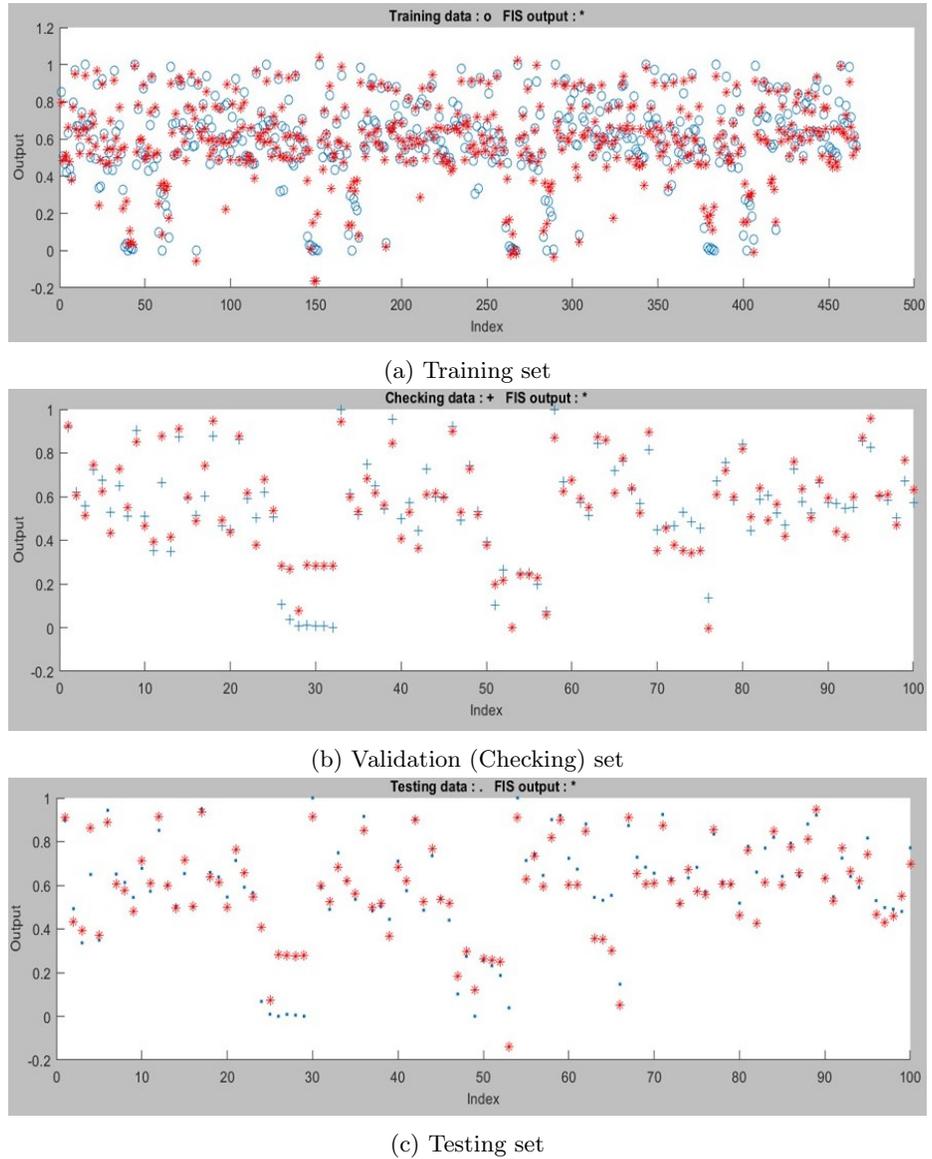

(a) Training set

(b) Validation (Checking) set

(c) Testing set

Fig. 11: Distribution of the predictions of the model and the original values over the three sets. Blue dots original data. Red stars: Model output

4.2   Comparison with other methods

Although the results show that the method is capable of performing the task with low error, it is necessary to compare its performance with other methods. There are no other previous works that address the issue of predicting the output power of a cell using EL images, so it is not possible to make a comparison with them. However, table 2 presents a comparison between the proposed model and some traditional machine learning methods implemented in Python using the Sklearn library. It can be seen that ANFIS has a significantly lower error than Linear Regression and Support Vector Machine. In the case of Gradient Boosting, both methods have similar performance, with Gradient Boosting slightly better in terms of MAE and ANFIS better in terms of MSE and RMSE. Even with similar results, ANFIS provides more robust results as it provides the set of rules that were used to make the inference.

Table 2: Comparison of the different traditional machine learning models with the proposed ANFIS method. MAE: Mean Absolute Error, MSE: Mean Squared Error, RSME: Root Square Mean Error

Metric	ANFIS	Linear Regressor	Support Vector Regressor	Gradient Boosting Regressor
MAE	0.0647	0.09169	0.06830	0.05933
MSE	0.00940	0.02153	0.01291	0.01058
RSME	0.09695	0.14673	0.11362	0.10285

Despite its results, the ANFIS system presents some drawbacks that need to be taken into account. This system is not ready at this moment for real-time applications since it only works at cell level. It also needs to be tested with other kinds of PV cells to determine and improve its generalization capacities. Another possible drawback is that its training is more computer-demanding than traditional Machine-learning but it is still much faster than Deep Learning methods.

5   Conclusions and future work

Predicting the power generated by PV cells is a complex problem. The model needs to be trained with data from the IV curve of the cell. This data is not easy to obtain as it is not common to obtain the IV curve of a single PV cell. In an attempt to develop a model capable of predicting the output power, a combination of Artificial Neural Networks and Fuzzy Logic is presented. This model has the learning capabilities of the ANN and the human-readable knowledge of Fuzzy Logic.

The resulting model is shown to be able to solve the problem more effectively than other traditional machine learning models, with an RSME of 0.09695. The model also produces 40 human-understandable rules that can be used to analyze

the behavior of the model or to find new patterns unknown to domain experts. However, the model is not without shortcomings, as it has only been trained on one type of PV cell (polycrystalline). Another improvement should be the further optimization of the membership functions, using methods such as genetic algorithms or other metaheuristics to optimize not only the number of functions for each input but also their shapes.

Finally, it would be extremely interesting to adapt the algorithm to be able to work in Real-Time applications but it would need a method to take the EL of the panels without harming the production of the PV installations and an optimized segmenting algorithm to divide the panels into cells.

6   Acknowledgements

References

1. Acciani, G., Falcone, O., Vergura, S.: Typical defects of pv-cells. In: 2010 IEEE International Symposium on Industrial Electronics. pp. 2745–2749 (2010). https://doi.org/10.1109/ISIE.2010.5636901
2. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L.: Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. J Big Data **8**, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8, https://doi.org/10.1186/s40537-021-00444-8
3. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L.: Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. J Big Data **8**, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8, https://doi.org/10.1186/s40537-021-00444-8
4. Chawla, R., Singal, P., Garg, A.K.: A mamdani fuzzy logic system to enhance solar cell micro-cracks image processing. 3D Research **9**, 1–12 (9 2018). https://doi.org/10.1007/S13319-018-0186-7/METRICS, https://link.springer.com/article/10.1007/s13319-018-0186-7
5. Hernández-Callejo, L., Gallardo-Saavedra, S., Alonso-Gómez, V.: A review of photovoltaic systems: Design, operation and maintenance. Solar Energy **188**, 426–440 (aug 2019). https://doi.org/10.1016/j.solener.2019.06.017, https://linkinghub.elsevier.com/retrieve/pii/S0038092X19305912
6. Jahic, A., Konjic, T., Pihler, J., Jahic, A.: Photovoltaic power output forecasting with anfis. In: Mediterranean Conference on Power Generation, Transmission, Distribution and Energy Conversion (MedPower 2016). pp. 1–8 (2016). https://doi.org/10.1049/cp.2016.1056

7. Jang, J.S.: Anfis: adaptive-network-based fuzzy inference system. IEEE Transactions on Systems, Man, and Cybernetics **23**(3), 665–685 (1993). https://doi.org/10.1109/21.256541

8. Karaboga, D., Kaya, E.: Adaptive network based fuzzy inference system (anfis) training approaches: a comprehensive survey. ARTIFICIAL INTELLIGENCE RE-VIEW **52**(4), 2263–2293 (DEC 2019). https://doi.org/10.1007/s10462-017-9610-2

9. Khosrojerdi, F., Taheri, S., Cretu, A.M.: An adaptive neuro-fuzzy inference system-based mppt controller for photovoltaic arrays. In: 2016 IEEE Electrical Power and Energy Conference (EPEC). pp. 1–6 (2016). https://doi.org/10.1109/EPEC.2016.7771794

10. Òscar Lorente, Riera, I., Rana, A.: Image classification with classic and deep learning techniques (2021)

11. Mateo Romero, H.F., Gonzalez Rebollo, M.A., Cardenoso-Payo, V., Alonso Gomez, V., Redondo Plaza, A., Moyo, R.T., Hernandez-Callejo, L.: Applications of artificial intelligence to photovoltaic systems: A review. Applied Sciences **12**(19) (2022). https://doi.org/10.3390/app121910056, https://www.mdpi.com/2076-3417/12/19/10056

12. Mateo-Romero, H.F., Hernandez-Callejo, L., Rebollo, M.A.G., Cardeñoso-Payo, V., Gomez, V.A., Bello, H.J., Moyo, R.T., Aragones, J.I.M.: Synthetic dataset of electroluminescence images of photovoltaic cells by deep convolutional generative adversarial networks. Sustainability **15**(9), 7175 (Apr 2023). https://doi.org/10.3390/su15097175, http://dx.doi.org/10.3390/su15097175

13. Mellit, A., Kalogirou, S.A.: Artificial intelligence techniques for photovoltaic applications: A review. Progress in Energy and Combustion Science **34**(5), 574–632 (2008). https://doi.org/https://doi.org/10.1016/j.pecs.2008.01.001, https://www.sciencedirect.com/science/article/pii/S0360128508000026

14. Mellit, A., Kalogirou, S.A.: Anfis-based modelling for photovoltaic power supply system: A case study. Renewable Energy **36**(1), 250–258 (2011). https://doi.org/https://doi.org/10.1016/j.renene.2010.06.028, https://www.sciencedirect.com/science/article/pii/S0960148110002843

15. Morales-Aragonés, J.I., Gómez, V.A., Gallardo-Saavedra, S., Redondo-Plaza, A., Fernández-Martínez, D., Hernández-Callejo, L.: Low-cost three-quadrant single solar cell i-v tracer. Applied Sciences **12**(13) (2022). https://doi.org/10.3390/app12136623

16. Pillai, D.S., Blaabjerg, F., Rajasekar, N.: A comparative evaluation of advanced fault detection approaches for pv systems. IEEE Journal of Photovoltaics **9**(2), 513–527 (2019). https://doi.org/10.1109/JPHOTOV.2019.2892189

17. REN21: Renewables 2022 Global Status Report. REN21 (2022), https://www.ren21.net/

18. Takagi, T., Sugeno, M.: Fuzzy identification of systems and its applications to modeling and control. IEEE Transactions on Systems, Man, and Cybernetics **SMC-15**(1), 116–132 (1985). https://doi.org/10.1109/TSMC.1985.6313399