Shingle Cell IV Characterization Based on Spatially Resolved Host Cell Measurements

Philipp Kunze Matthias Demant Alexander Krieg Ammar Tummalieh Nico Woehrle Stefan Rein*

Philipp Kunze*
Fraunhofer Institute for Solar Energy Systems
Heidenhofstraße 2
79110 Freiburg im Breisgau
Germany
Email Address: philipp.kunze@ise.fraunhofer.de

Dr. Matthias Demant, Alexander Krieg, Ammar Tummalieh, Dr. Nico Woehrle, Dr. Stefan Rein
Fraunhofer Institute for Solar Energy Systems
Heidenhofstraße 2
79110 Freiburg im Breisgau
Germany
Email Addresses: firstname.lastname@ise.fraunhofer.de

Keywords: *Solar Cell, Shingle Cell, Photovoltaic, Machine Learning, Deep Learning, Characterization*

Each solar cell is characterized at the end-of-line using current-voltage (IV) measurements, except shingle cells, due to multiplied measurement efforts. Therefore, the respective host cell quality is adopted for all resulting shingles, which is sufficient for samples with laterally homogeneous quality. Yet, for heterogeneous defect distributions, this procedure leads to (i) loss of high quality shingles due to defects on neighboring host cell parts, (ii) increased mismatch losses due to inaccurate binning and (iii) lack of shingle-precise characterization. In spatially resolved host measurements, such as electroluminescence images, all shingles are visible along with their properties. Within a comprehensive experiment 840 hosts and their resulting shingles are measured. Thereafter, a deep learning model has been designed and optimized which processes host-images and determines IV parameters like efficiency or fill factor, IV curves and binning classes for each shingle cell. The efficiency can be determined with an error of $0.06\,\%_{\mathrm{abs}}$ enabling a $13\,\%_{\mathrm{abs}}$ improvement in correct assignment of shingles to bin classes compared to industry standard. This results in lower mismatch losses and higher output power on module level as demonstrated within simulations. Also IV curves of defective and defect-free shingle cells can be derived with good agreement to actual shingle measurements.

1 Introduction

At the end-of-line, solar cells are usually either sorted into bin classes based on their current-voltage (IV) characteristics or sorted out if they have severe defects. Yet, when it comes to shingle cells, those measurements are cumbersome making this approach unattractive for industry. Since several shingle cells are cut from one host cell, the measurement effort per Watt would be multiplied when measuring them individually. Hence, production speed would be considerably decreased. Therefore, the host cell is usually measured before separation. Presuming homogeneity, identical host quality is assumed for the shingle cells leading to same quality classes and characteristics for each of them.
Simply applying host characteristics to all resulting shingle cells leads to inaccurate characterization and subsequent issues as false positive rejections or reduced module quality. Our experiments described in this paper show Pearson correlation coefficients of only 60 - 70 % between host and shingle IV parameters. Accordingly, this can potentially lead to reduced sorting quality and hence higher mismatch losses, i.e. performance losses due to dissimilarity of assembled cells within a module. In addition, higher material loss can be expected. In our experiments, about 40 % of the shingle cells were incorrectly rejected as they are situated on a locally faulty host.
The potential inhomogeneity in the processing of the host wafer is, however, well visible in spatially resolved inline measurements. Figure 1 shows on the left the electroluminescence (EL) image of one host

cell. The upper shingles exhibit scratches; additionally, point-like inhomogeneities can be seen in the area of the lowermost shingle. If only the host IV characteristic is considered for sorting, all six shingle cells would be discarded. On the right, the efficiency and bin class of each shingle cell is shown according to the shingle measurements. The efficiencies are slightly reduced compared to the host cell due to the cutting process which leads to slightly increased edge recombination. Yet, it can be seen that only three of the six shingles would have to be sorted out, while the other three could be integrated into modules. One shingle cell even is among the best of the dataset studied. These characteristics are well visible in the EL image of the host cell and can be used when characterizing the shingles.

We present a deep learning approach in which a convolutional neural network (CNN) is designed and optimized to derive the quality of the resulting shingle cells from host measurements only. For this purpose, we created a dataset consisting of comprehensive host and shingle measurements before and after separation, respectively. We use industry standard inline measurements such as EL, photoluminescence (PL) as well as thermography (TG) measurements of the host cell. The models are designed to evaluate the individual shingle regions and approximate their quality, resulting in six outputs for each host cell, one per shingle. We develop several models to study the prediction quality of single IV parameters such as efficiency or fill factor, as well as the prediction quality for whole IV curves for each shingle cell. We evaluate our approach in terms of its effect on module level by simulating and comparing the mismatch loss of the different sorting methods.

In other works, it has been shown that CNNs are well suited to incorporate spatially resolved inline measurements into cell characterization. There exist various approaches to derive quality parameters of cells from EL or PL images [1–11]. Those, however, are still focusing on full cells only. Also for modules, CNNs in combination with EL or TG images are employed for quality assurance. Meanwhile, defect detection in measured images using machine learning methods is becoming a common procedure for cells and wafers as well, as indicated by the number of publications regarding that [12–33]. In terms of subcell characterization, Buratti *et al.* have presented an attempt to characterize half-cells via stretching the half image to full cell size [34, 35]. In contrast, we developed a compact model using full non-stretched host-images and providing six shingle quality parameters or six shingle IV curves at once.

Specifically, our contributions are (i) the development of empirical machine learning models combining host measurements to determine shingle quality, (ii) a thorough evaluation of these models in terms of IV parameters, binning quality, and IV curves with respect to real measurements and reference models, and (iii) module simulations to investigate the sorting schemes' influence in terms of output power and mismatch loss.

2 Approach

We propose a deep learning model to compute IV parameters and IV curves of shingle cells from their host measurements. The structure of the model is shown schematically in Figure 1 for the use case of efficiency η prediction and binning. As shown in this work, the model can be trained for the prediction of other IV parameters and IV curves as well. The input measurements of the host cell are shown on the left: We utilize EL, TG and PL images, additionally the respective host IV data is used as input. In the use case illustrated, the shingle efficiencies are predicted, therefore the host efficiency η_{host} is passed as input. For shingle evaluation, this data is processed by a CNN. This results in feature maps as shown in the middle part under Shingle Extraction. From these, the corresponding shingle regions are filtered out as highlighted in white. Regarding the example cell, it is visible in the top feature map that the defects in the top two shingles have been identified. This becomes clear from the lighter areas in the feature map, which represent higher activation in these areas and appear to be screening for this type of defect. For the respective shingle regions, we apply shingle average pooling which also excludes the pseudo-squares, resulting in one feature vector per shingle. These are each linearly mapped to the target variables, in this case η_1 - η_6 for the six shingles, respectively.

Furthermore, Figure 1 shows the corresponding bin class of the host cell and the shingles, which are partially different. The host efficiency of $\eta_{host} = 21.3\,\%$ corresponds to a low value within the considered

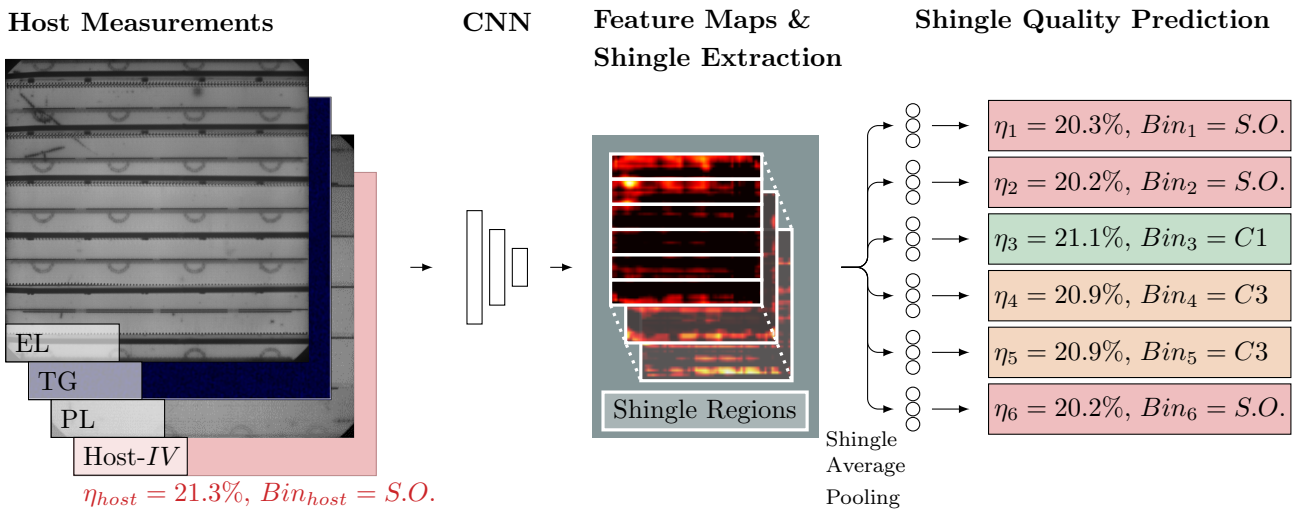


Figure 1: Structure of the empirical model for the calculation of shingle quality. Electroluminescence (EL), termography (TG), photoluminescence (PL) and host IV measurements are shown on the left. These are passed to a convolutional neural network (CNN) that produces feature maps reflecting properties of the shingles. Lighter areas represent higher activations of the corresponding feature map. The individual shingle regions are pooled and within linear mappings used to predict the IV parameters, IV curves, and bin classes of the shingles.

dataset. According to the low η_{host}, all shingle cells would be sorted out ($Bin_{host} = S.O.$). The cutting process slightly reduces the efficiency of the shingles due to an increased edge recombination. The criterion for "sorted out" (S.O.) is an efficiency $\eta_{shingle} < 20.7\,\%$. Furthermore, we define five bin classes C5 - C1 in $0.1\,\%_{\text{abs}}$ steps upwards, C5 being the worst and C1 the best class. When examining the shingles' efficiencies, it can be seen that only three of the six shingles would need to be sorted out while the remaining three get assigned to bin classes C1 or C3, respectively.
For derivation of shingle IV curves, a current vector $\mathbf{i}$ is predicted for 100 specified voltage values. The fixed voltage values were not set to be uniformly distributed, but were rather positioned to more densely map key regions around the short-circuit current (J_{sc}), maximum power point (MPP), and open-circuit voltage (V_{oc}). Also, as suggested in another work [1], the loss function was extended to the first and second derivatives as described in Equation (1).

$$\mathrm{L} = \mathrm{L}_1(\hat{\mathbf{i}}, \mathbf{i}) + \mathrm{L}_1(\hat{\dot{\mathbf{i}}}_{1-40}, \dot{\mathbf{i}}_{1-40}) + \mathrm{L}_1(\hat{\ddot{\mathbf{i}}}_{1-40}, \ddot{\mathbf{i}}_{1-40}) \tag{1}$$

L_1 denotes the L1 loss function, $\hat{\mathbf{i}}$ is the predicted current vector, $\mathbf{i}$ is the measured shingle current vector. The dots stand for the respective derivative. The indices stand for the first 40 entries of the vectors, which concern approximately the range from J_{sc} to MPP. The inclusion of gradient terms in our loss function favors a smooth IV progression between neighboring points that follows the measurement curve.

3 Experimental

As a basis for the experiments, we use 620 monofacial industrially produced passivated emitter and rear cells (PERCs) as hosts. In addition, we use 220 host PERCs that were rejected by the manufacturer due to optical or electrical defects in order to achieve a large variance in the dataset. The hosts are of M6 format, i.e. of size $166 \times 166\,\mathrm{mm}^2$ with pseudo-square meaning edge and middle shingle vary slightly in area. Ga-doped p-type silicon as base material was used to ensure stable cell efficiencies at room-temperature. Each of the host cells was tagged six times with a unique data matrix code in each of the shingle regions. The host cells were measured using EL, PL, TG and IV. For the EL measurements a measurement device from halm. GmbH with an excitation current of 10 A, an integration time of 200 ms and a gain of 3 was used. IV and TG measurements were also performed by means of this device at reverse conditions using $-12\,\mathrm{V}$. Each of the still-connected shingles within the host was centrally contacted on the front and rear

side with a contact bar. A measuring device from ISRA Vision was used for the PL images.
Each of the host cells was separated into six shingles using the laser scribe and mechanical cleave (LSMC) method [36–38]. Starting with 840 host cells, this results in roughly 5000 shingle cells. The LSMC process requires manual splitting, which may result in greater shingle variation compared to an automated industrial process. Each of the shingle cells was measured by the same measuring equipment used for the host cells. The shingles were centrally contacted with contact bars for the IV measurements. The unique identifier on each shingle was then used to build a dataset in which each of the six shingles can be assigned to the corresponding host cell.
For the models described in Section 2, we used a modification of DenseNet [39]. The models were implemented in PyTorch [40] and the Adam Optimizer [41] was utilized with a batch size of 20 and a learning rate of 10^{-4} and a plateau learning rate reduction schedule. All optimizations were performed on a Nvidia GeForce RTX 2080 Ti graphics card. The models were trained for 100 epochs and evaluated within a 5-fold cross-validation. Accordingly, four fifths of the dataset were used to optimize each model and tested on the last fifth. The process was repeated five times, each time using a different fifth as the test set, so that prediction results are available for each cell without it being in the training set.

3.1 Experiment 1: Shingle IV Parameter Estimation

We investigate how well the image-based models can predict the IV parameters of the shingle cells from the host measurements. For this purpose, we specifically analyze the prediction of the efficiency η, the fill factor FF, the open-circuit voltage V_{oc} and the short-circuit current density J_{sc}. Since PL measurements are not always available in industry, we additionally examine to which extent the results change when the model is optimized without PL images of the hosts. Furthermore, we study the individual errors with respect to the shingles' position within the host cell. As quality criteria, we use the mean absolute error (MAE) and the Pearson correlation coefficient (ρ).
For reference, the host IV measurements are utilized with mean edge losses due to recombination at the non-passivated shingle edges being subtracted. E.g. regarding the efficiency, we calculate the mean efficiency difference $\Delta\overline{\eta}$ between host efficiency $\overline{\eta}_{hosts}$ and shingle efficiency $\overline{\eta}_{shingles}$ as defined in Equation (2).

$$\Delta\overline{\eta} = \overline{\eta}_{hosts} - \overline{\eta}_{shingles} \tag{2}$$

Afterwards, this value is subtracted from the efficiency of the current host cell η_{host} in the reference method to obtain the reference efficiency of the shingle cells, as described in Equation (3).

$$\eta_{shingle,ref} = \eta_{host} - \Delta\overline{\eta} \tag{3}$$

The other IV parameters are calculated accordingly.

3.2 Experiment 2: Binning

In this experiment, we study the extent to which binning quality can be improved using the image-based models compared to the industry standard. The binning is therefore performed via the image-based model using the predicted efficiency and the binning classes as described in Section 2. We compare the accuracy of bin sorting and examine each binning class individually.
In addition, we investigate the impact of using an image-based model for binning at module level. For this, we sort the shingles (i) based on the shingle measurements as reference, (ii) using the empirical image-based model and (iii) based on the host IV measurements. Shingles sorted out by the respective method are not considered, since they would not be built into a module. Within a module simulation, the hypothetical module performance is calculated under standard test conditions (STC). The modules consist of six strings connected in parallel. Each string contains 60 shingles connected in series. To get more consistent results, we copy the shingles not being sorted out by the respective sorting scheme eight times and randomize the the order of shingle cell selection within each bin class via the respective sorting approach.
For quantitative evaluation, we evaluate the power output of the respective classes and sorting approaches.

For this, we compare the output powers of the three sorting schemes per bin class for all modules that are simulated. We also consider the number of modules that can be build with respect to the sorting schemes. Additionally, we use the mismatch loss P_{ml} as evaluation criterion as defined in Equation (4).

$$P_{ml} = P_{sum} - P_{sim} \tag{4}$$

P_{sum} corresponds to the sum of the power output at MPP of the individual shingle cells interconnected within one module and P_{sim} to the module power according to the module simulations. The difference P_{ml} is considered as mismatch loss coming from shingle cells' dissimilarity.

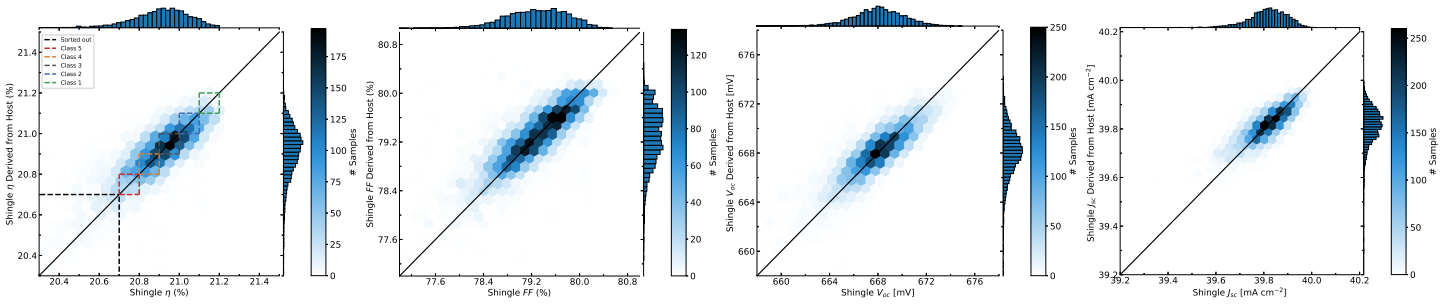
3.3 Experiment 3: Shingle IV Curve Estimation

Within this experiment, we investigate how well the shingle IV curves can be derived from the host measurement images. For this purpose, using the approach described in Section 2, the IV curves of some example shingles with and without defects will be examined and compared with the actual IV curve measured. A comparison to the host IV curve is considered as well. Additionally, the mean absolute error and correlation coefficient per fixed voltage value are calculated over the dataset.

4 Experimental Results

4.1 Experiment 1: Shingle IV Parameter Estimation - Results

Correlation plots for all IV parameters visualize the agreement between predicted and measured quantities. In Figure 2, the IV parameters estimated by the model from the host measurements (y-axis) are plotted against the actual measurements (x-axis) in the form of a hexa-bin histogram. In accordance with the colorbar on the right, darker bins correspond to areas with many points. The results are visualized in Figure 2 (a) for η, in Figure 2 (b) for FF, in Figure 2 (c) for V_{oc} and in Figure 2 (d) for J_{sc}. Ideally, all points should lie on the black diagonal ($x = y$). For all parameters examined, the predictions cluster close to the diagonal. The results for η in Figure 2 (a) and V_{oc} in 2 (c) seem to be particularly accurate, as the dark hexabins corresponding to the respective colorbar are located on or very close to the diagonal.
The results of our investigations show, that the prediction quality does not change strongly when the model is trained with or without PL image as additional input. Table 1 distinguishes the mean absolute errors (MAE) and correlation coefficients (ρ) for the IV parameters of the shingles once with and once without PL measurement in columns 2-3 and 4-5, respectively. The IV parameters can be predicted with and without PL image with errors lower than the typical measurement uncertainty and high correlation coefficient. For η and FF the correlation coefficients are higher with PL image than without, the MAEs are comparable, for FF the error is even slightly lower without PL image. Regarding V_{oc} and J_{sc} there are almost no differences.
Our model prediction correlates to the shingle measurements and outperforms the reference method purely based on host IV. In addition to the results of the CNN, the results of the host based method can be seen in columns 6-7. The errors are higher than those of the empirical models. The correlation coefficient increases by more than $20\,\%_{\mathrm{abs}}$ for η and FF, for example, and J_{sc} also shows more than a halving of the MAE. With regard to V_{oc}, there is an increase of $10\,\%_{\mathrm{abs}}$ for the correlation coefficient, the mean errors, however, are only slightly improved.
The IV parameters of the shingle cells can be predicted at all positions within the host, although slightly increased errors occur at the edge shingles. In Table 2, the errors of the investigated IV parameters (column 2-5) are displayed as a function of the shingle position (column 1). Positions 1 and 6 correspond to the edge shingles, the remaining positions to inner shingles within the host cell. For all parameters considered, it can be seen that the errors at individual shingle level are similar to the global mean values from Table 1. However, for the edge shingle cells the errors of all parameters are slightly increased compared to those of the inner shingle cells. In particular, this can be seen for J_{sc} (column 5), where the two edge shingles both show an error increased by $0.02\,\mathrm{mA\ cm^{-2}}$ compared to positions 2-5 whose background is discussed in Section 5.



(a) (b) (c) (d)

Figure 2: Comparison of the predicted (y-axis) and measured (x-axis) IV parameters of the shingle cells. In (a) the results are shown with respect to η, in (b) with respect to fill factor FF, in (c) with respect to open-circuit voltage V_{oc}, and in (d) with respect to short-circuit current density J_{sc}.

Table 1: Results concerning the IV parameter derivation of the shingle cells. Columns 2-5 show the mean absolute errors and correlation coefficients of the empirical model once with and once without PL images as model input. Columns 6-7 display the results based on only host IV measurements as a reference (reflecting the present standard).

	Image-Based with PL		Image-Based without PL		Host-Based	
Parameter	MAE	ρ	MAE	ρ	MAE	ρ
η	0.06 $\%_{\text{abs}}$	82.9 %	0.06 $\%_{\text{abs}}$	75.4 %	0.11 $\%_{\text{abs}}$	60.0 %
FF	0.21 $\%_{\text{abs}}$	82.5 %	0.13 $\%_{\text{abs}}$	60.0 %	0.33 $\%_{\text{abs}}$	60.4 %
V_{oc}	0.92 mV	87.9 %	0.92 mV	87.8 %	1.10 mV	77.6 %
J_{sc}	0.03 mA cm^{-2}	71.4 %	0.05 mA cm^{-2}	73.0 %	0.07 mA cm^{-2}	66.4 %

4.2 Experiment 2: Binning - Results

The binning of the shingle cells can be improved by our network model based on image measurements by about 13.6 $\%_{\text{abs}}$ on average. In Figure 3, the distribution of true shingle classes according the IV measurements at the separated shingle cells is depicted for each classification bin when assigned by two different prediction models. The respective left bar stands for the host-based model using only the IV data of the host cells. The right bars stand for the image-based empirical model using the IV and image data of the host cells. Green colored parts correspond to shingle cells, which have been assigned to the correct class by the respective method. According to the legend, the other colors stand for the actually correct bin class based on the shingle IV measurement. As can be seen, the image-based approach increases the proportion of correctly assigned shingles for all bin classes. Also, a larger proportion of misclassified shingles are in adjacent bin classes compared to the host-based method. On average, the sorting accuracy is increased by about 13.6 $\%_{\text{abs}}$ meaning that the amount of shingles which are classified to their actual correct bin is increased by this percentage with the image-based model compared to the host-based approach. This is reflected by the increase of the green bars in Figure 3.

Based on the various binning approaches, we want to assemble homogeneous modules with the same cell quality. The output power of the simulated modules between shingle-measurement-based and image-

Table 2: Overview of mean absolute errors of IV parameter results for shingle cells depending on their position within the host cell.

Shingle	Mean Absolute Errors			
Position	η [$\%_{abs}$]	FF [$\%_{abs}$]	V_{oc} [mV]	J_{sc} [mA cm^{-2}]
1	0.07	0.24	0.94	0.05
2	0.06	0.22	0.88	0.03
3	0.05	0.21	0.85	0.03
4	0.05	0.22	0.90	0.03
5	0.06	0.21	0.90	0.03
6	0.07	0.27	1.04	0.05

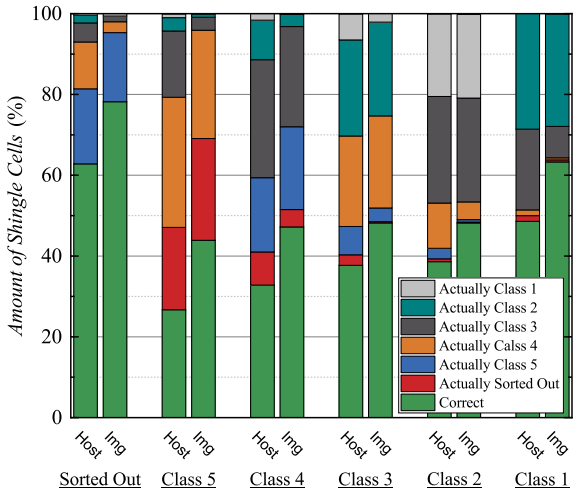


Figure 3: Binning accuracy with respect to the sorting classes using the reference method based on only IV data of the host cells (left bars) and the our model based on IV and EL and PL image data of the host cells (right bars).

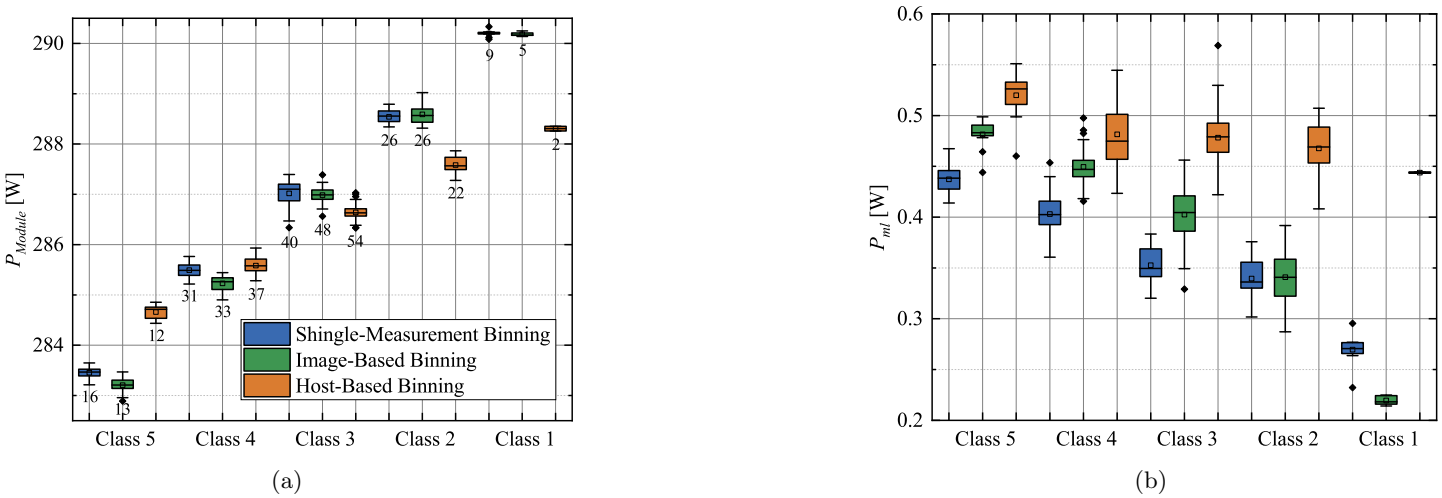
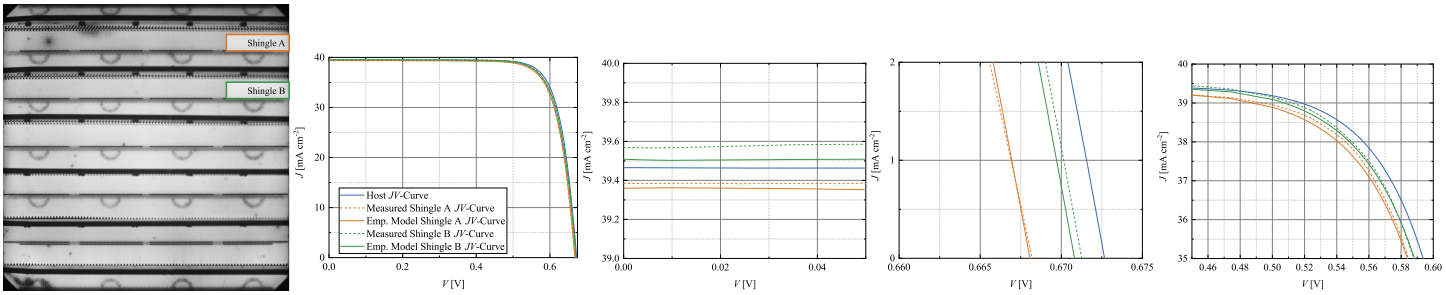


Figure 4: (a) Output power and (b) missmatch losses of the simulated modules as a function of the bin classes for the different sorting schemes (colors).

based approach is similar, while the host-based approach causes the modules to cover a narrower range over the output power. Figure 4 (a) depicts the simulated power outputs of the modules as a function of the bin classes as a box plot. The blue boxes represent the shingle-measurement-based approach as a reference, while the green ones represent the image-based and the orange ones the host-based approach. The numbers below each box indicate the number of modules assigned to the corresponding bin class by the respective sorting scheme. The image-based and the reference approach have similar performance per bin class. The number of assigned modules is also similar. The largest variation here is found in C3 with 40 and 48 modules, respectively. The module performance seems to increase by a similar amount per bin class. When host-based sorting is considered, it can be seen that the bin class distribution is more dense, so that the power classes are close to each other. With this method, most modules are assigned to the middle bin classes C4 and C3, and the number of better-class modules in C1 and C2 is lower than for the image-based and reference sorting schemes. Also, the average performance of C1 modules is ca. 2 Watt lower than for the reference sorting, while the average performance of the C5 modules is increased by ca. 1 Watt.

In Figure 4 (b), the mismatch with respect to all bin classes are shown in blue for the reference, in green for the image-based approach, and in orange for the host-based approach. Both the reference and the image-based approach show slightly lower mismatch losses than the host-based approach. The losses of the latter are also approximately constant for all bin classes, while those of the reference and image-based



(a) (b) (c) (d) (e)

Figure 5: Visualization of shingle JV curves predicted from host measurements for an example cell. (a) EL image of the host cell with shingles A (orange) and B (green) whose JV-curves are further evaluated. (b-d) JV curve predictions (continuous) and measurements (dashed) for shingles A and B in comparison to the host JV curve (blue) (b) as a whole and zoomed into the regions around (c) J_{sc}, (d) V_{oc} and (e) MPP.

approaches decrease for better bin classes. The modules of class C1 have even lower mismatch losses with the image-based approach than those sorted with shingle measurements.

4.3 Experiment 3: Shingle IV Curve Estimation - Results

The IV curves of the shingles calculated by the image-based model from the host measurements match well with the measured curves. In Figure 5 an example host cell and the IV curves results can be seen. In Figure 5 (a) the EL image of the host cell is depicted. Shingle A marked in orange has some point-like defects (dark spots) possibly originiating from a local contamination during processing, which lowers its quality. Shingle cell B marked in green does not show any visible defects. In Figure 5 (b-d), the current density J over the voltage V is displayed for these two shingles in the corresponding colors. The complete curves are shown in Figure 5 (b) and details in the regions around J_{sc}, V_{oc} and MPP in Figure 5 (c-e), respectively. The dashed curves in each figure represent the measurements, while the solid ones show the prediction of the model. In addition, the JV curve of the host cell is depicted in blue. The measured and predicted curves of shingle A lie close to each other and exhibit lower J_{sc} and V_{oc} values due to the present point defect. A slightly different slope compared to the measurement can be observed in the region around V_{oc} in Figure 5 (d) along with a slight deviation around MPP in Figure 5 (e). The defect-free shingle B is also well approximated. In the region around J_{sc}, it can be seen that the model-based JV-curve is lying below the measurement by about 0.06 mA cm^{-2}, at V_{oc} the curves run parallel at a distance of about 0.7 mV. In the region around MPP the lines also lie close to each other. Compared to the host measurement, the approximation of the model is slightly more accurate, however, the model slightly underestimates the current values in these examples.

If defects are not visible in the measurement images, the prediction accuracy of the image-based model is reduced. Figure 6 shows EL images and IV curves of another host cell. In Figure 6 (a), the bar-contacted EL image is shown, which is also input to the image-based model. In Figure 6 (b), the EL image of the same cell with a GridTouch contacting unit is shown. Shingle cell B is highlighted in red. It can be seen that shingle B has a long scratch that is hidden by the contacting bar in Figure 6 (a). In Figure 6 (c), the continuous line is the predicted JV curve and the dashed one is the measured JV curve of the shingle. It can be seen that they do not lie on top of each other. In Figure 6 (d) an excerpt of the curve around the MPP is shown. It is visible that the measured curve is below the predicted curve of the shingle cell, while the JV curve of the host cell is even above both shingle curves which reflects the positive impact of the other defect-free shingle cells.

On average, the prediction of the voltage-dependent current seems to be more accurate in the range of I_{sc} than in the range of V_{oc}. In Figure 7, the Pearson correlation coefficient ρ and the mean absolute error between shingle prediction and measurement are shown in orange and blue, respectively. In the range from 0 to 0.5 V, there is a low error of about 0.0002 A and a high correlation coefficient of about 95 %. From there to 0.675 V the correlation coefficient decreases and the error increases. The mean V_{oc} of the dataset

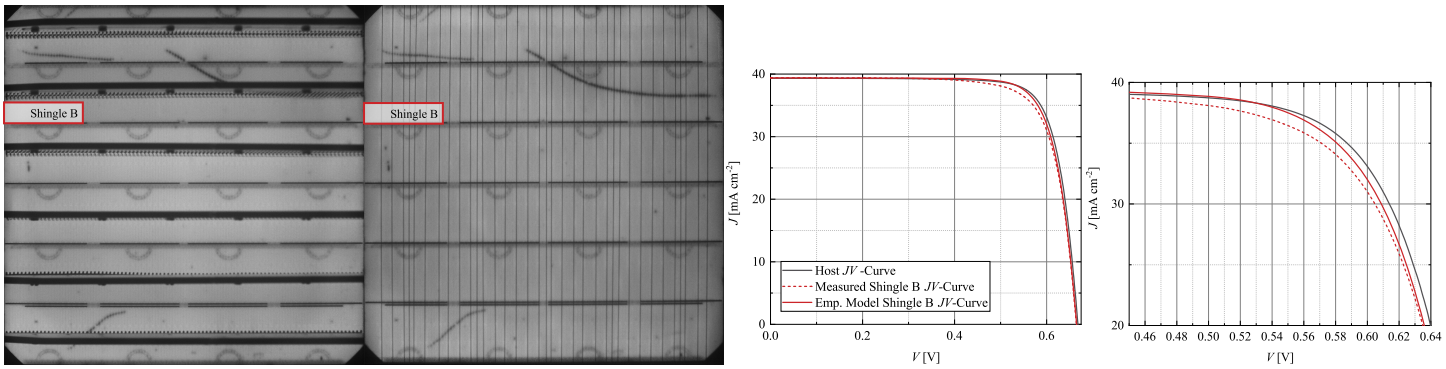


Figure 6: Visualization of the JV curve calculation of a second example cell. In (a) and (b) EL images of the same host cell are shown, measured with a bar contact unit and the wire-based GridTouch contact unit, respectively. The investigated shingle cell B is highlighted in red. The predicted (red, continuous) and measured (red, dashed) shingle JV curves and the host JV curves are displayed in (c) and (d).

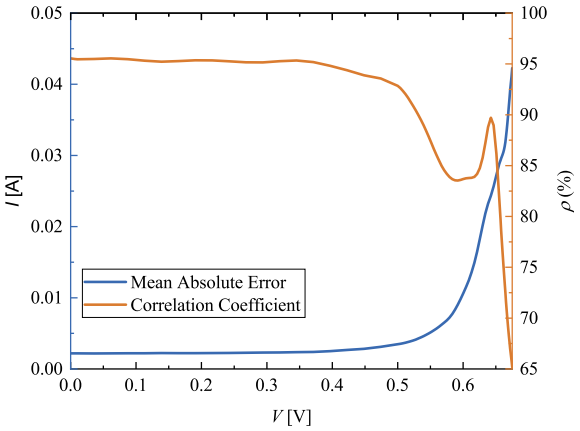


Figure 7: Visualization of mean absolute error (blue) and correlation factor (orange) of measured and predicted current I as a function of voltage V for shingle IV curve prediction.

is about 0.667 V, at which the current prediction error is 0.034 A. Due to the steep slope in this region, larger errors are to be expected.

5 Discussion

The prediction of common IV parameters of shingle cells from spatially resolved host inline measurements is possible with reasonable accuracy. For the parameters studied, the results of the empirical models are better than those of the host-IV-based method in terms of mean absolute error and correlation coefficients. The method also works without the inclusion of PL images. The results of the networks deteriorate only slightly in this case. Surprisingly, the error becomes smaller in the case of FF (Table 1), but the correlation coefficient also decreases by 20 $\%_{\text{abs}}$. It seems that the edge shingles are harder to be assessed by the model since the error decreases for the center shingles. An explanation is that more cell area is obscured by the contacting bars due to the angle of the camera for the edge shingles, which is well visible in Figures 5 (a) and 6 (a) compared to the middle shingles. The occluded view increases the uncertainty of the empirical model within those regions and results in higher prediction errors.
The image-based approach is well suited for sorting shingle cells into bin classes and improves the output results on module level. The sorting results are more accurate compared to the host-based approach. Shingles being incorrectly sorted by the empirical approach are mostly located in neighboring bin classes, while the classification error sorting distribution of the host-based classification is wider. That results in more frequent modules built up e.g. from shingles coming from bin classes C3 and C1 and by this reducing the

overall module quality. Accordingly, the performance groups of the image-based approach are separated more effectively compared to the host-based approach, so that the performance of the image-binned modules reaches above 290 W and that of the host-based ones is just at ca. 288 W. Furthermore, fewer modules of the best class were generally produced by host-binning than by image-binning within our investigations. However, it should be noted that instead of treating the shingles equally as described in Equation (3), the reference method could be improved by handling center and edge shingles separately.

The model is also capable of deriving the IV curves of the shingles from spatially resolved measurements of the hosts, but is limited in some cases by occluded cell properties. The JV curves of the shingles are well reproduced and are in better agreement withe the actual shingle cell measurement than the host JV curve. This is the case for both, defective and defect-free shingles. However, it is possible that a defect may be obscured by measurement equipment such as the bar contact in the EL measurement, so that the hidden defect will cause discrepancies in the prediction. In these cases, the model does not appear to be piecing together the concealed information from the other measurements, such as the PL image. The problem could be reduced by using more data in which defects are partially occluded or by selectively covering known defects while training enforcing the model to stitch together partly occluded defects. Furthermore, it would be thinkable to use only PL images, possibly in combination with a partial shading [1], in order to measure series resistance properties as well. In addition, a GridTouch contact unit could be used covering less cell area. Disadvantage here is that it reveals series resistance defects less effectively. The error in the prediction increases in the region between MPP and V_{oc}. This is expected, since in this region, the largest change in the IV curve is present. Furthermore, the EL measurement is performed approximately at J_{sc} conditions with 10 A, but the defects arising from separation mainly affect FF and V_{oc}. The losses due to separation cannot be derived from the host measurements which may result in a larger uncertainty in the regions around MPP and V_{oc}.

We assess the methodology to be relevant for today's shingle cell production, however, as it involves little additional effort in practice. Since the empirical model is based on electroluminescence, photoluminescence, thermography and IV measurements of the host, which are widely used measurement techniques in industry, no additional overhead is required here. The evaluation of the measurements can happen at inline speed. However, an additional step of sorting the shingles is necessary, which is not required using the reference method based on only host IV data, since the shingles from each host cell are sorted into the same quality class irrespective of local variations. In addition, the approach is based on the assumption that the losses due to the cutting process are approximately constant. A systematic defect in this process step would not be detected by the model. Nevertheless, this also applies to the current industry standard.

6 Conclusion

We have introduced a deep learning model to characterize shingle solar cells as an extension to the existing idea of adopting host IV values. The approach is based on common spatially resolved inline measurement techniques such as electroluminescence or photoluminescence measurements of the host cells that have not yet been cut. On these measured images, the properties of the shingles are visible and can be assessed within a convolutional neural network. We have developed models regarding IV parameters, binning and IV curve prediction.

The empirical models have shown to be able to derive IV parameters, IV curves and binning classes of the shingles from host measurements accurately. For example, mean absolute errors with respect to the efficiency η of 0.06 $\%_{\text{abs}}$ or with respect to short-circuit current density J_{sc} of 0.03 mA cm^{-2} are obtained with high correlation coefficients. These values are higher than those of the reference method based purely on host IV values. Accordingly, binning can be improved by 13 $\%_{\text{abs}}$ which results in clearly separable module classes and reduced mismatch loss. Also IV curves of the shingles can be well derived from the host measurements for both defective and defect-free shingles.

We expect that better binning will result in less wasted shingle cells because fewer cells will be incorrectly rejected. In addition, the quality of the shingles can be assessed during the process without any loss of throughput. Lastly, less mismatch loss in the module can be anticipated due to more accurate binning.

Supporting Information
Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements
This work was partly funded by the German Federal Ministry for Economic Affairs and Climate Action within the project BiZePS-Plus (03EE1064). P. Kunze would like to thank the Studienstiftung des deutschen Volkes for the support within the scope of his scholarship.

References

[1] P. Kunze, J. M. Greulich, A. Tummalieh, W. Wirtz, H. Hoeffler, N. Woehrle, S. Glunz, S. Rein, M. Demant, *Solar RRL* **2022**, 2200599.

[2] A. M. Karimi, J. S. Fada, N. A. Parrilla, B. G. Pierce, M. Koyutürk, R. H. French, J. L. Braid, *IEEE Journal of Photovoltaics* **2020**, *10*, 3 878.

[3] A. Kovvali, M. Demant, T. Trötschler, J. Haunschild, S. Rein, In *AIP Conference Proceedings*, volume 1999. AIP Publishing LLC, **2018** 130011.

[4] M. Demant, P. Virtue, A. Kovvali, X. Y. Stella, S. Rein, *IEEE Journal of Photovoltaics* **2019**, *9*, 4 1073.

[5] M. Demant, P. Virtue, A. Kovvali, X. Y. Stella, S. Rein, *IEEE Journal of Photovoltaics* **2019**, *9*, 4 1064.

[6] A. S. Kovvali, M. Demant, S. Rein, In *2019 IEEE 46th Photovoltaic Specialists Conference (PVSC)*. IEEE, **2019** 3059–3062.

[7] M. Hoffmann, C. Buerhop-Lutz, L. Reeb, T. Pickel, T. Winkler, B. Doll, T. Würfl, I. Marius Peters, C. Brabec, A. Maier, et al., *Progress in Photovoltaics: Research and Applications* **2021**, *29*, 8 920.

[8] P. Kunze, S. Rein, M. Hemsendorf, K. Ramspeck, M. Demant, *Solar RRL* **2022**, *6*, 5 2100483.

[9] M. Demant, L. Kurumundayil, P. Kunze, A. Woernhoer, A. Kovvali, S. Rein **2020**.

[10] S. Rodrigues Abreu, J. Hauch, I. M. Peters, T. Winkler, C. Buerhop-Lutz, C. Brabec, B. Doll, In *36th PV Symposium*, FZJ-2021-02794. Helmholtz-Institut Erlangen-Nürnberg Erneuerbare Energien, **2021** 97–110.

[11] Y. Buratti, A. Sowmya, R. Dumbrell, P. Dwivedi, T. Trupke, Z. Hameiri, *Joule* **2022**.

[12] B. Su, H. Chen, Z. Zhou, *IEEE Transactions on Industrial Electronics* **2021**, *69*, 3 3161.

[13] R. Pierdicca, E. Malinverni, F. Piccinini, M. Paolanti, A. Felicetti, P. Zingaretti, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* **2018**, *42*, 2.

[14] Y. Jiang, C. Zhao, W. Ding, L. Hong, Q. Shen, In *2020 IEEE 9th Data Driven Control and Learning Systems Conference (DDCLS)*. IEEE, **2020** 1415–1421.

[15] J. Balzategui, L. Eciolaza, N. Arana-Arexolaleiba, J. Altube, J.-P. Aguerre, I. Legarda-Ereño, A. Apraiz, In *2019 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, **2019** 529–535.

[16] A. Bartler, L. Mauch, B. Yang, M. Reuter, L. Stoicescu, In *2018 26th European signal processing conference (EUSIPCO)*. IEEE, **2018** 2035–2039.

[17] Z. Ying, M. Li, W. Tong, C. Haiyong, In *2018 Chinese Automation Congress (CAC)*. IEEE, **2019** 3571–3576.

[18] O. D. Singh, A. Malik, V. Yadav, S. Gupta, S. Dora, *Multimedia Tools and Applications* **2021**, *80*, 5 6509.

[19] B. Su, H. Chen, P. Chen, G. Bian, K. Liu, W. Liu, *IEEE Transactions on Industrial Informatics* **2020**, *17*, 6 4084.

[20] S. Deitsch, V. Christlein, S. Berger, C. Buerhop-Lutz, A. Maier, F. Gallwitz, C. Riess, *Solar Energy* **2019**, *185* 455.

[21] M. Mayr, M. Hoffmann, A. Maier, V. Christlein, In *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, **2019** 1885–1889.

[22] H. Chen, S. Wang, J. Xing, In *2019 Chinese Automation Congress (CAC)*. IEEE, **2019** 2565–2569.

[23] Y. Zhao, K. Zhan, Z. Wang, W. Shen, *Progress in Photovoltaics: Research and Applications* **2021**, *29*, 4 471.

[24] A. Ahmad, Y. Jin, C. Zhu, I. Javed, A. Maqsood, M. W. Akram, *IET Renewable Power Generation* **2020**, *14*, 14 2693.

[25] A. Chindarkkar, S. Priyadarshi, N. S. Shiradkar, A. Kottantharayil, R. Velmurugan, In *2020 47th IEEE Photovoltaic Specialists Conference (PVSC)*. IEEE, **2020** 1612–1616.

[26] E. Sovetkin, E. J. Achterberg, T. Weber, B. E. Pieters, *IEEE Journal of Photovoltaics* **2020**, *11*, 2 444.

[27] M. W. Akram, G. Li, Y. Jin, X. Chen, C. Zhu, X. Zhao, A. Khaliq, M. Faheem, A. Ahmad, *Energy* **2019**, *189* 116319.

[28] M. R. U. Rahman, H. Chen, *IEEE Access* **2020**, *8* 40547.

[29] X. Zhang, T. Hou, Y. Hao, H. Shangguan, A. Wang, S. Peng, *IEEE Access* **2021**, *9* 62093.

[30] A. V. de Oliveira, M. Aghaei, R. Rüther, In *Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition, Marseille, France*. **2019** 9–13.

[31] L. Liu, Y. Zhu, M. R. U. Rahman, P. Zhao, H. Chen, In *2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI)*. IEEE, **2019** 292–297.

[32] M. Demant, T. Welschehold, S. Kluska, S. Rein, *IEEE Journal of Photovoltaics* **2015**, *6*, 1 136.

[33] P. Kunze, J. Greulich, S. Rein, K. Ramspeck, M. Hemsendorf, A. Vetter, M. Demant, In *Presented at the 37th European PV Solar Energy Conference and Exhibition*, volume 7. **2020** 11.

[34] Y. Buratti, A. Sowmya, R. Evans, T. Trupke, Z. Hameiri, In *2020 47th IEEE Photovoltaic Specialists Conference (PVSC)*. IEEE, **2020** 0133–0138.

[35] Y. Buratti, A. Sowmya, R. Evans, T. Trupke, Z. Hameiri, *Progress in Photovoltaics: Research and Applications* **2022**, *30*, 3 276.

[36] P. Baliozian, E. Lohmüller, T. Fellmeth, N. Wöhrle, A. Krieg, R. Preu, *Solar RRL* **2018**, *2*, 3 1700171.

[37] P. Baliozian, M. Al-Akash, E. Lohmüller, A. Richter, T. Fellmeth, A. Münzer, N. Wöhrle, P. Saint-Cast, H. Stolzenburg, A. Spribille, et al., *IEEE Journal of Photovoltaics* **2020**, *10*, 2 390.

[38] P. Baliozian, E. Lohmüller, T. Fellmeth, A. Richter, A. Münzer, A. Bhandary, N. Wöhrle, A. Spribille, R. Preu, In *2020 47th IEEE Photovoltaic Specialists Conference (PVSC)*. IEEE, **2020** 0476–0480.

[39] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, In *Proceedings of the IEEE conference on computer vision and pattern recognition*. **2017** 4700–4708.

[40] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, et al., *Advances in neural information processing systems* **2019**, *32*.

[41] D. P. Kingma, J. Ba, *arXiv preprint arXiv:1412.6980* **2014**.