

Visualizing Material Quality and Similarity of mc-Si Wafers Learned by Convolutional Regression Networks

Matthias Demant , Patrick Virtue, Aditya Kovvali, Stella X. Yu, and Stefan Rein

Abstract—Convolutional neural networks can be trained to assess the material quality of multicrystalline silicon wafers. A successful rating model has been presented in a related work, which directly evaluates the photoluminescence (PL) image of the wafer to predict the current–voltage parameters after solar cell production. This paper presents the results of two visualization techniques to understand what has been learned in the network. First, we reveal what has been learned in the PL image by visualizing the spatial quality distribution of the wafers based on the activation maps of the network. The method is denoted as regression activation mapping. We compare regression activation maps with j_0 images of solar cells to show the semantically meaningful representation of the trained features. Second, we show what has been learned in the data by mapping the learned network representation of all wafers into a low-dimensional subspace. Visualizations reveal the smoothness of our representation with respect to the PL input and measured quality. This technique can be used to detect material anomalies or process faults for samples with high prediction errors.

Index Terms—Convolutional neural network (CNN), densely connected convolutional neural network (denseNet), fault detection, machine learning, material, multicrystalline silicon (mc-Si), passivated emitter and rear cell, photoluminescence (PL), rating, regression, regression activation mapping, solar cell, t-distributed stochastic neighborhood embedding (t-SNE).

I. INTRODUCTION

CONVOLUTIONAL neural networks (CNNs) can be used to rate the quality of multicrystalline silicon (mc-Si) wafers based on photoluminescence (PL) images [1] with high

Manuscript received June 24, 2018; revised November 26, 2018; accepted February 15, 2019. This work was supported by the German Federal Ministry for Economic Affairs and Energy through the Project “Q-Crystal” under Contract 0324103A. The work of M. Demant was supported by a fellowship of the German Academic Exchange Service within the program FITweltweit. (Corresponding author: Matthias Demant.)

M. Demant is with the Fraunhofer Institute for Solar Energy, 79100 Freiburg, Germany, and also with the International Computer Science Institute, University of California, Berkeley, CA 94704 USA (e-mail: matthias.demant@ise.fraunhofer.de).

P. Virtue and S. X. Yu are with the International Computer Science Institute, University of California, Berkeley, CA 94704 USA (e-mail: virtue@eecs.berkeley.edu; stellayu@berkeley.edu).

A. Kovvali and S. Rein are with the Fraunhofer Institute for Solar Energy, 79100 Freiburg, Germany (e-mail: aditya.kovvali@ise.fraunhofer.de; stefan.rein@ise.fraunhofer.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JPHOTOV.2019.2906037

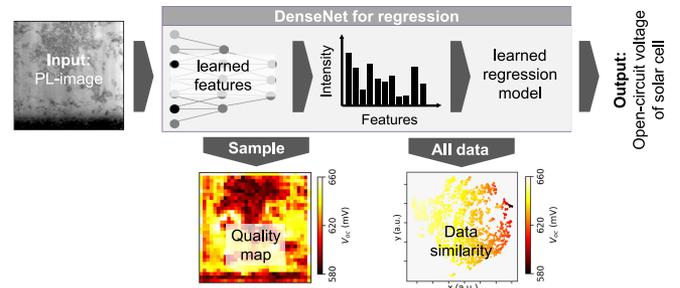


Fig. 1. We introduce an end-to-end rating model, which connects PL images as input and open-circuit voltages V_{oc} as output (see Section II and [1]). The network representation will be used by two visualization techniques to understand what has been learned: 1) a quality map for each wafer shows the learned quality distribution in low resolution; and 2) a 2-D map reveals the learned similarity of samples in the dataset.

prediction accuracy [2]. A novelty of this approach is the end-to-end connection between input image and quality output: a CNN can learn a direct mapping between the PL image of the wafer and the quality of the resulting solar cell based on empirical data only. In contrast with traditional feature engineering approaches, as proposed, e.g., in [3]–[6], a network does not rely on a description of the wafers with human-designed features. Yet, CNN-based efficiency predictions deviate from the measure efficiency in average about 35%_{rel} less than the predictions with our previous feature-engineering approach, as presented in the first part of this study [1].

Deep learning allows the discovery of relevant features, but requires additional steps for understanding what has been learned. The features trained from a deep CNN are highly complex. On the one hand, the learned features are purely based on the empirical training data. This helps us avoid errors due to handcrafted features, which may not be capable to address all variations of data. On the other hand, this information is spread along the layers of the network, and the contribution of an image region or image feature is unclear.

This paper evaluates two techniques for visualizing, what has been learned by the regression network for open-circuit voltage prediction (V_{oc}), as shown in Fig. 1. A quality map reveals what has been learned in the image (see Sections III and V-D), and an embedding technique shows which similarities have been learned in the dataset (see Sections IV and V-E).

What has been learned in the image? Visualization techniques increase the interpretability of what has been learned within the

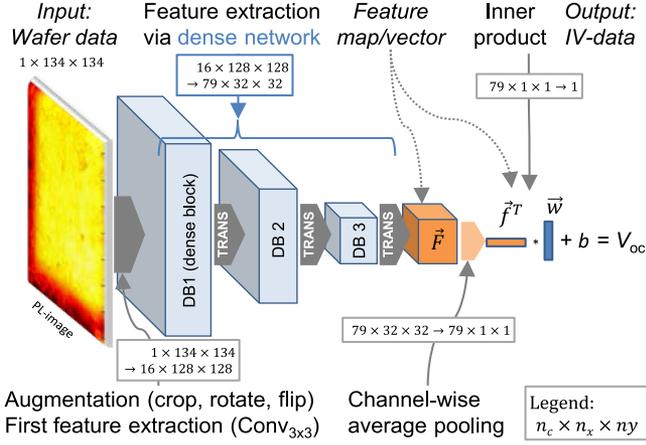


Fig. 2. Scheme of the convolutional regression network for V_{oc} prediction based on PL images of as-cut wafers. The network contains three densely connected (DB) blocks and transition layers (TRANS). The feature map \vec{F} and feature vector \vec{f} of a sample are used for data visualization.

models. These methods have been presented, for example, by Zeiler and Fergus [7] and Yosinski *et al.* [8]. In most classification problems, a specific image region is responsible for the decision. The most decisive image regions for a classification are revealed by mapping the activation of deep features back to input space. Zhou *et al.* [9] propose to localize these regions by analyzing the activation of the network in the final layers. Within this work, we apply the localization method to regression tasks. Therefore, we rearrange the final feature layers and weights to form an image of spatially resolved quality values, which is denoted as regression activation map (RAM) [10]. The average value of the map is equal to the regression result.

What has been learned in the dataset? The features in the last layers of our network are a semantic representation of the image, which can be used to analyze similarity and continuity in our data with respect to PL input and V_{oc} output. Similar defect distributions in the PL images are expected to lead to similar feature values and similar open-circuit voltages. Embedding techniques like the t-distributed stochastic neighborhood embedding (t-SNE) [11] allow the structure of high-dimensional data to be visualized in a lower dimensional space. Input images and quality data can be compared within the map, as shown, e.g., in [12]. We analyze this human-interpretable representation of the data to see what has been learned.

II. DENSELY CONNECTED CONVOLUTIONAL NEURAL NETWORKS FOR REGRESSION

In the following, we briefly introduce our network design depicted in Fig. 2. The network is based on a densely connected convolutional neural network (DenseNet) introduced by Huang *et al.* [13]. Since the network was designed for classification problems, it has to be transformed into a regression network to predict the V_{oc} based on the input PL images. Further details on

the feature extraction via the dense network are visualized in [1] and described in the following.

In the first step, the PL image is downsampled to 134×134 pixels, due to memory constraints of the GPUs used in this experiment. The augmentation step is an optional step, which transforms the image to avoid an overfitting of the data to the training set: the number of rows and columns are randomly cropped by 6 pixels, and the image is randomly rotated or flipped.

A convolutional layer with filter size 3×3 ($\text{Conv}_{3 \times 3}$) extracts 16 feature channels. This feature map with size $16 \times 128 \times 128$ is the input for the dense network.

The dense network extracts image features within three densely connected blocks. Each block contains sequences of convolutional layers for feature extraction, batch normalization layers to scale the data, and rectifying linear units as nonlinear operations. Skip connections from each layer to all succeeding layers in a dense block allow the reuse of the features and a robust training of the model.

Transition layers connect the dense blocks. The spatial resolution of the features is reduced by maximum pooling for the first two transition layers, which increases the receptive field of each neuron. The feature channels are compressed by 1×1 convolutions, which combines the feature channels in a meaningful way with regard to V_{oc} prediction.

The dense network results in a feature map \vec{F}_i with 79 feature channels c and spatial resolution in x and y of 32×32 pixels for each sample i . A feature vector \vec{f}_i is extracted by channel-wise average pooling along the elements n_{px} of each channel according to

$$\vec{f}_i(c) = \frac{1}{n_{px}} \sum_{x,y} \vec{F}_i(c, x, y). \quad (1)$$

The feature vector $\vec{f}_i \in \mathbb{R}^{79}$ is used to predict the open-circuit voltage \hat{y}_i within an inner product layer with weights \vec{w} and offset b resulting with

$$\hat{y}_i = \langle \vec{w}, \vec{f}_i \rangle + b. \quad (2)$$

The network is optimized during training with stochastic gradient descent minimizing the loss between the predicted quality value \hat{y}_i and the measured y_i . For the test case, we average the prediction result for nine different cropping positions of the sample.

III. DISCOVERING THE SPATIAL DISTRIBUTION OF THE PREDICTED QUALITY IN THE REGRESSION ACTIVATION MAP

We apply the localization method of Zhou *et al.* [9] to regression networks to visualize the spatial distribution of the predicted quality value \hat{y}_i as an image with the same average value. The method is applicable for regression networks, where the final layer contains a channelwise averaging of the feature map and a subsequent inner-product layer. The predicted value \hat{y}_i can be computed as pixelwise average along all channels c of the

weighted feature map by reformulating (1) and (2):

$$\hat{y}_i = \sum_c \vec{w}(c) \cdot \frac{1}{n_{\text{px}}} \sum_{x,y} \vec{F}_i(c, x, y) + b \quad (3)$$

$$= \frac{1}{n_{\text{px}}} \sum_{x,y} \underbrace{\left(\sum_c \vec{w}(c) \cdot \vec{F}_i(c, x, y) + b \right)}_{\vec{Y}_i(\text{RAM})}. \quad (4)$$

Each feature channel of the feature map $\vec{F}_i(c, x, y)$ is weighted by the corresponding coefficient $\vec{w}(c)$ of the inner-product layer. The pixelwise averaging along all feature channels leads to a RAM \vec{Y}_i with the same spatial resolution as the feature map.

IV. UNDERSTANDING THE DATA IN THE LEARNED FEATURE SPACE

A human-interpretable visualization of the learned feature space can be created to understand the learned similarities of the data. The network reduces the data in the PL image to a lower number of features \vec{f}_i in the final layer. These features presumably contain a semantic meaningful representation of the input data with respect to the corresponding quality values. The feature dimension, here $n_c = 79$, is still too high for a human interpretation of what has been learned. Therefore, the data are visualized via a t-SNE [11] to analyze the learned network representation.

The t-SNE approach maps each data point onto a typically two-dimensional (2-D) manifold and preserves the local structure of these high-dimensional data: Very similar data points in the high-dimensional space are likely to be neighbored in the lower dimensional map. The dimension reduction algorithm minimizes an objective function that measures the discrepancy between local similarities in the high-dimensional data and similarities in the map. The dimensions of the t-SNE map have arbitrary units.

The t-SNE map allows a human observation of what has been learned: We can visualize the similarities between different data in the deep feature space. Therefore, PL images are passed through the network. The high-dimensional features are extracted for each sample and mapped into a 2-D space via t-SNE. The mapping is computed for all data based on the final feature vector \vec{f}_i after average pooling. Each data point in the t-SNE map corresponds to a sample with a measured PL image and V_{oc} . PL images can be visualized in a ‘‘Patchwork’’ image for selected samples at the corresponding position in the t-SNE map. In addition, the V_{oc} values can be visualized according to the t-SNE coordinates.

‘‘Patchwork’’ image: An image is created, which visualizes the input PL images at the position of the data point in the t-SNE map with low resolution. Not all of the data can be visualized within one image. Therefore, the 2-D t-SNE map is sampled following a regular grid with fixed step size. For each grid point, we are searching for the nearest sample in the map within the region of patch. The input image of this sample will be added to the patchwork at this position in a low resolution.

Quality distribution: The distribution of the V_{oc} is visualized within a scatter plot in the t-SNE map. Regions of high and low material quality are compared with the input images at the corresponding locations in the patchwork image.

V. EXPERIMENTAL DETAILS

A. Dataset

The dataset contains 7300 mc-Si wafers and high-performance multicrystalline silicon (HPMC-Si) [14] wafers with a size of 156×156 mm from 74 bricks and 25 boxes of ten different manufacturers. The as-cut wafers were measured with an inline PL system with a line-scan camera with an InGaAs detector, as described by Höffler *et al.* [15] with a resolution of 1024×1024 pixels. The samples are processed to passivated emitter and rear cells [16] within an industrial production line. The current–voltage characteristics are measured after solar cell production. For simplicity, the trained model has been optimized for V_{oc} prediction only.

1) *Scenario 1—Testing ‘‘Unknown’’ Bricks:* For the evaluation of the activation map, a model is trained with about 2900 wafers from selected 32 bricks and seven boxes and a validation set of 425 wafers for parameter optimization. The model is tested with about 3900 wafers from distinct 42 bricks and 18 boxes not represented in the training set.

2) *Scenario 2—Testing ‘‘Unknown’’ Manufacturers:* We investigate if a rating model can be used to evaluate new materials, i.e., HPMC-Si wafers, from an ‘‘unknown’’ manufacturer with an existing prediction model. This is relevant if we want to rate the potential of new materials for a given solar cell production based on PL images only. In the past, new material classes like HPMC-Si [14], [17] have been developed, which can lead to higher solar cell efficiency compared with previous mc-Si materials. The prediction of ‘‘unknown’’ materials is challenging especially for this high-quality materials, since the measurement of bulk lifetime in PL images of as-cut wafers is limited by surface recombination.

Therefore, we evaluate the materials from the manufacturers with HPMC-Si wafers M1, M2, and M3 separately. We create three datasets by removing all wafers from a manufacturer from the training set in Scenario 1 and add these samples to the test set.

B. Implementation and Training

The network is implemented and evaluated using the deep learning framework Caffe [18]. The input images are aligned [19, pp. 58–59], downsampled, and normalized before the training process. The training was accomplished via stochastic gradient descent with a batch size of 20 samples using Nesterov optimization for 400k iterations. Following a graduated scheme, the learning rate was initialized to 0.01 and halved after periods of 50k iterations.

C. Results on Quality Prediction

The network predicts the V_{oc} for materials from ‘‘unknown’’ bricks (Scenario 1) with low mean absolute prediction errors

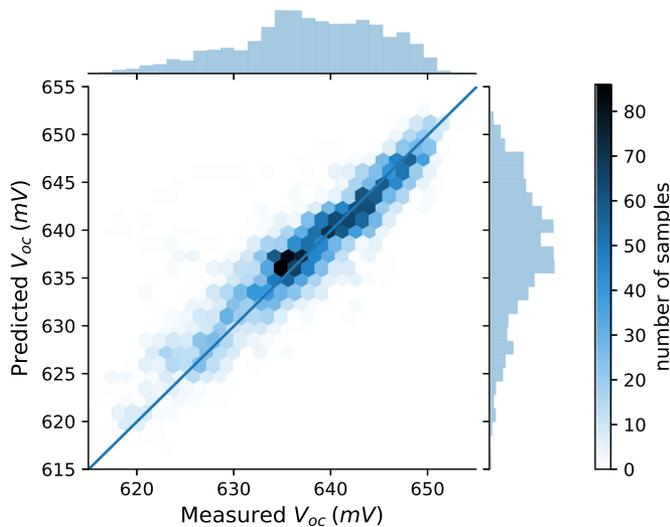


Fig. 3. Measured and predicted V_{oc} values for samples from “unknown” bricks (Scenario 1). The color of each hexagon indicates the number of samples within this region. In addition, the margin distribution is shown for both parameters.

TABLE I
DENSENET V_{oc} PREDICTION FOR SCENARIO 2 (“UNKNOWN” MANUFACTURER)

| Test data | Ingots # | Bricks # | Wafers # | mae [mV] | rms [mV] | cor [-1,1] |
|-----------|-------------|-------------|-------------|-------------|-------------|---------------|
| M1 | 2 | 8 | 685 | 1.87 | 2.5 | 0.91 |
| M2 | 1 | 4 | 182 | 1.64 | 2.19 | 0.94 |
| M3* | 1 | 6 | 434 | 5.09 | 8.37 | 0.84 |

mae: mean absolute error, rms: root mean squared error,
cor: Pearson correlation coefficient.

*full testset will be used for anomaly detection due to high errors.

of 2.12 mV and the Pearson correlation coefficient of 0.93, as shown in Fig. 3. Further results of an extended multivariate regression network predicting four current–voltage parameters using PL images with different resolutions are presented in [1].

The prediction errors for unknown HPMC-Si manufacturers (Scenario 2) are low for materials from M1 and M2, but high for materials from M3, as shown in Table I. The prediction errors are an indicator that manufacturer M3 has samples with anomalies according to the learned patterns of our network. These are referable to process errors or material defects not visible in band-to-band PL images, as discussed in Section IV.

D. Results on Activation Mapping

The presented activation map is implicitly learned by the DenseNet and represents the predicted quality value with spatial distribution. Therefore, a sample is passed forward through the network. Then, the RAM is computed according to the activation in the final feature map of the sample and the inner-product coefficients of the network (see Section III). The maps are given in millivolts with the same spatial resolution of 32×32 pixels as the final feature map.

We compare triplets of representative PL images, RAMs, and images of the dark saturation current j_0 . Fig. 4 shows examples of computed maps (second column) from given PL images (first column) and models for V_{oc} prediction. The PL images are

the input for the neural network, which means that they have been cropped and normalized. The measured V_{oc} is shown below the PL image. The average value of the RAM corresponds to the predicted V_{oc} for this PL image and is given for each map. The j_0 image of the corresponding solar cell (third column) is determined according to [20] and represents the spatial distribution of the defects, which are actually recombination active in the finished solar cell and, thus, limiting V_{oc} . In addition, the j_0 images are cropped. All images show the same region of the wafer and cell with different spatial resolution.

The network learned that specific structures are bad. Without previous user input, the network learned to assign a low V_{oc} value to contaminated regions and dislocation structures. The prediction prefers the structure of the defects over the pure intensity, as shown, e.g., in Fig. 4(a) and (b): Although PL intensities in contaminated regions can be lower than PL intensities of structural defects, they are not necessarily assigned to lower V_{oc} values. This is in accordance with observations in the j_0 images, as the lowest PL intensities observed in contaminated regions do not necessarily correspond to the highest j_0 values [e.g., Fig. 4(c)].

Low-quality structures in the RAMs are similar to structures of low intensity in the PL image and high intensity in j_0 images, but do not exactly match the same pattern. The position and shape of the quality deterioration in the RAM can deviate from the exact defect position in PL images due to the size of the receptive field at a pixel position. Each neuron activation is connected to a combination of convolution and pooling steps. Although the kernel size is small for each convolution, a sequence of convolutions increases the receptive field of a neuron. Nevertheless, a high similarity can be observed in almost every example.

The RAM cannot anticipate the impact of the solar cell process with high spatial accuracy. In particular, high-temperature processes during the emitter diffusion and contact formation can change the distribution of regions with reduced lifetime and, thus, the appearance of j_0 images compared with the PL images. For example, a gettering effect for contaminated regions can be observed in the j_0 image of Fig. 4(c). This effect cannot be observed with high spatial accuracy in the RAM. Nevertheless, it can be considered to be part of the overall weighting, since prediction errors are low for most of the samples with contaminated regions in Fig. 4.

Up to a certain extent, PL imaging and surface artifacts are considered as nonrelevant by the network. The PL image in Fig. 4(d) shows an edge wafer with vertical stripes in the PL image, due to saw marks or heterogeneous illumination. Most of these line structures do not lead to strong quality deterioration in the RAM. One vertical structure can be observed in the RAM, which may be connected to a strong measurement or surface artifact.

Even for a complicated PL structure with superimposed defects, the network outputs RAMs in the correct quality range. Fig. 4(e) and (f) shows examples of PL images with inverted PL contrast in regions of low bulk lifetime, where grain boundaries appear as bright instead of dark structures. The open-circuit voltages of these samples differ more than 20 mV. Fig. 4(e) shows the high-quality sample, whose quality is only slightly

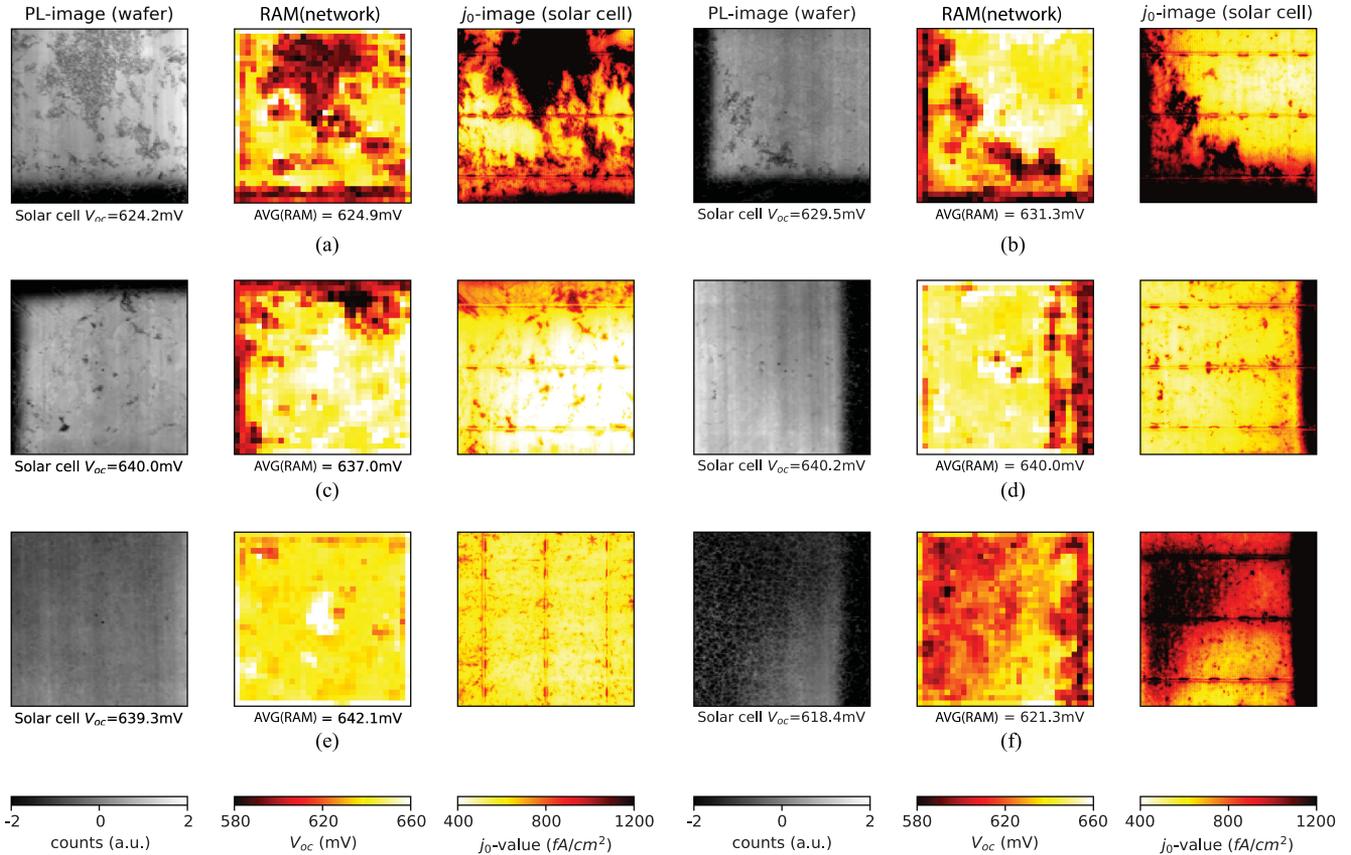


Fig. 4. Triplets of (left) the PL image of the as-cut wafer, (center) RAM learned by the network for V_{oc} prediction, and (right) the j_0 image [20] measured at the finished cells. The measured and predicted V_{oc} values are annotated below the left and center image, respectively. All images within a column are scaled equally, and all images show the same central image region according to the PL image region used by the network.

overestimated by the model. The RAM reveals a mostly homogeneous high-quality assignment of the model along the wafer. Fig. 4(f) shows a low-quality wafer. Again, the PL structures are rated similar to the quality shown in the j_0 image.

E. Results on Understanding the Data in Feature Space

The network representation of the PL images is a key to understand the data from a network perspective. The network learns to describe and compress the image data by means of quality-relevant defect features, which are represented in the last feature vector for each quality parameter. By visualizing the learned representation for our dataset in a 2-D space, we analyze the smoothness of the model and visualize erroneous samples.

1) *Dataset and Model*: The visualization technique can be used for error analysis. Therefore, we analyze the model with the most faulty prediction. The presented data are based on a model trained according to Scenario 2 (“unknown manufacturer”) without samples from manufacturer M3. Following the approach in Section IV, the deep features are extracted, and a 2-D embedding is learned via t-SNE.

We visualize the similarity of material defects and material quality given by the deep features. Similar material types according to the deep feature space are expected to be placed on similar locations in the 2-D manifold. The distribution of the

measured and predicted wafer quality, as well as the prediction error, is shown in Fig. 5. A human accessible visualization of the input images along the map is shown in the “Patchwork” image in Fig. 6.

2) *Material Quality and Photoluminescence Images*: The measured material quality varies smoothly along the t-SNE coordinates, as shown in Fig. 5 (left). Samples represented in the upper left region have a high material quality, and low-quality samples are located in the quadrants on the right. The t-SNE maps with measured V_{oc} in Fig. 5 (left) and predicted V_{oc} in Fig. 5 (center) show high similarity, besides few outliers in Fig. 5 (right). The comparison with the PL images in the “Patchwork” image in Fig. 6 shows that very low quality wafers are top and bottom wafers from various manufacturers. Materials in the upper left quadrant are HPMC-Si materials. Furthermore, low-quality clusters in the t-SNE map may be explained by either samples with strong dislocations or contaminated regions, which are located in similar regions of the map. For contaminated regions, the orientation of the samples becomes obvious in Fig. 6. The change in orientation for neighboring samples indicates that the network representation captures the semantic defect information independent of sample rotation.

3) *Error Analysis*: We can use the network errors for anomaly detection. PL images from M3 are highlighted in the Patchwork image shown in Fig. 6. The green and red color

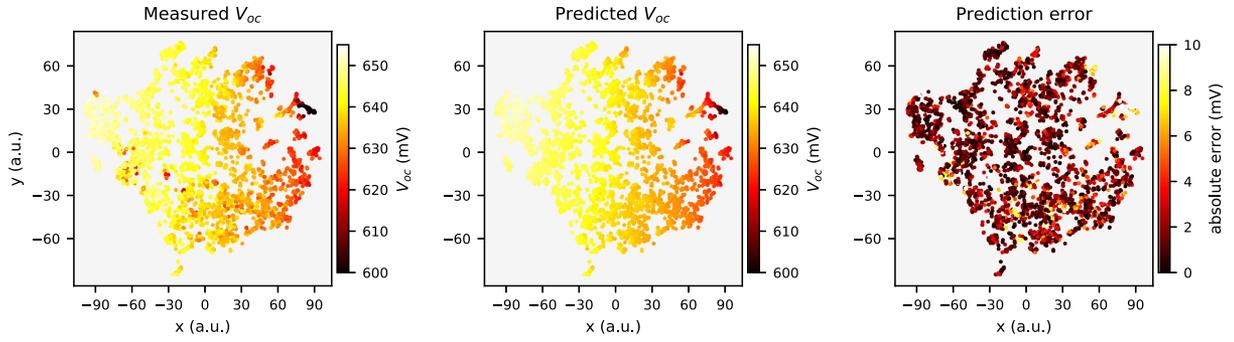


Fig. 5. Visualization of the high-dimensional deep features for all test data within a structure-preserving 2-D t-SNE map: (left) measured V_{oc} , (middle) predicted V_{oc} , and (right) the V_{oc} prediction error of the test data. The model was trained with data from Scenario 2 without samples from manufacturer M3.

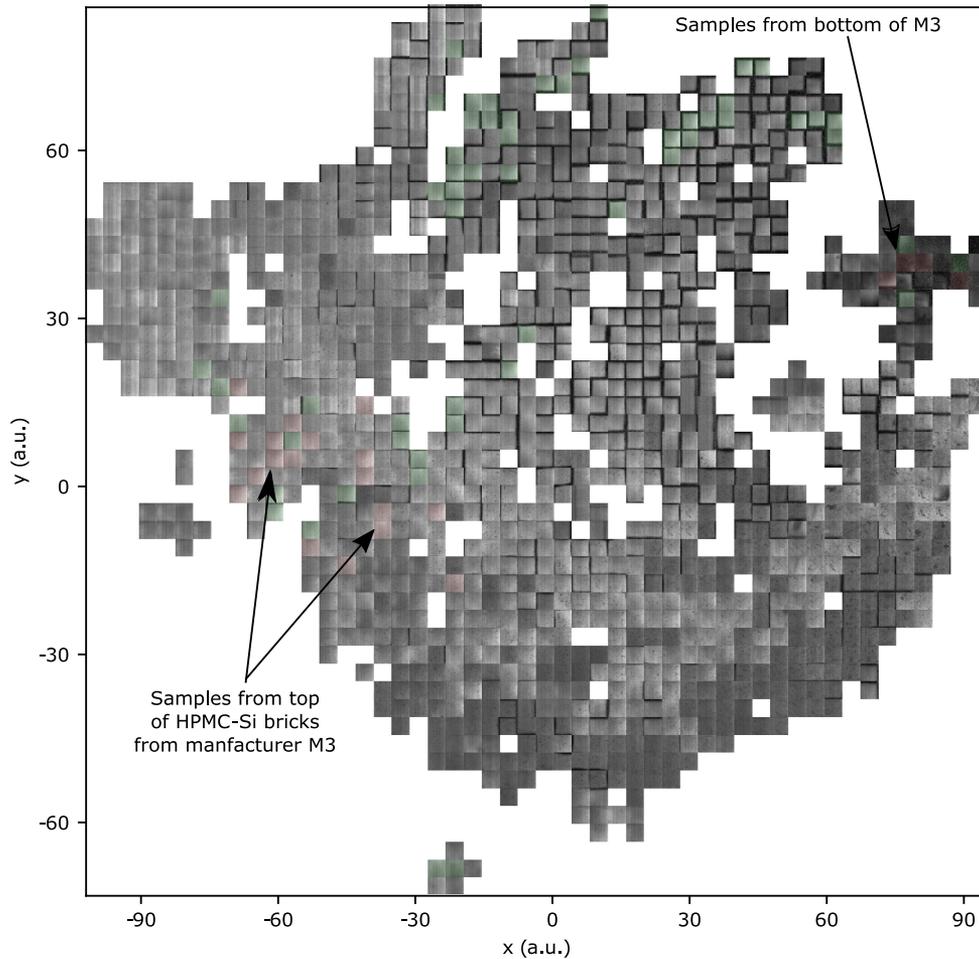


Fig. 6. Visualization of PL images according to the arrangement of the deep features of the sample after mapping them with a structure preserving projection onto a 2-D space (same as in Fig. 5). Each patch in the “Patchwork” image shows a PL image with reduced resolution at the corresponding position of the sample in the t-SNE map. The colored patches are HPMC-Si samples from manufacturer M3. Green and red colors indicate samples with prediction errors smaller and larger 6 mV, respectively. Prediction errors can be used for fault or anomaly detection. (Zoom in for details.)

overlay indicates samples with errors smaller and greater than 6 mV, respectively. Samples in the top region of bricks from M3 have smaller x -coordinates within the t-SNE map than the bottom wafers, which are located in the very right region of the t-SNE map. We can observe that top wafers from M3 lead to large prediction errors despite a high similarity to neighboring

wafers. Solely, a slight increase in dislocations can be observed in PL images from the top region.

Based on this fault detection, the samples are investigated in more detail. As the impurity concentration is typically higher in the top region of brick, a back diffusion of impurities from the very top of the brick is a possible cause for the observed

errors. Therefore, the concentration of interstitial iron point defects [21] has been determined in [1], and an increased number of metastable defects were measured for these samples. Although an increased concentration of point defects lowers the material quality in the top region of the bricks, the appearance of the PL images did not change significantly.

VI. DISCUSSION AND OUTLOOK

A. Discussion and Refinement of Regression Activation Maps

The activation mapping gives an insight into the “black box” neural network. This can improve the acceptance of network techniques in the photovoltaic community.

The activation map shows that the network has learned to assign reasonable quality values to most of the patterns with regard to a purely empirical approach based on a limited number of PL images. Without human input, dislocations and contaminated regions have been rated as a low-quality structure. Even low- and high-quality samples with superimposing patterns can be distinguished. Yet, the absolute quality values need to be analyzed in more detail.

Further investigations can be conducted to evaluate: 1) the robustness of the activation maps to image and surface distortions; 2) the impact of solar cell processes; and 3) material properties invisible in PL images. Therefore, activation maps for materials with different surface roughness or bulk defect distribution can be compared considering different solar cell processes. Further characterization with spatially resolved quality images of wafers and solar cells [22] allows a more quantitative evaluation of the mapping technique.

B. Understanding What Has Been Learned By Feature Space

Mapping the compressed feature representation of the final network layer onto a 2-D map allows a human interpretation of the data.

A smooth representation is an indicator for the generalizability of the model. The defect characteristics in the PL images and the quality data vary smoothly along the map. They are not exclusively clustered in groups of samples from the same brick. The smooth representation of the test data is an indicator that our representation does not underlie an overfitting to the training data.

The mapping can be applied for quality control: a crystal grower can monitor the different process results by comparing the map positions of new and previous processes. In the presented approach, failure cases for V_{oc} prediction are identified, which are referred to impurities in this material, as visualized in [1].

VII. CONCLUSION

In this paper, a machine learning approach for material rating of HPMC-Si and mc-Si wafers for solar cell production has been introduced. A CNN was trained to predict the I - V parameters based on PL images of as-cut wafers.

Two visualization techniques has been proposed to allow a human interpretation of the rating models: By applying the regression activation mapping, the rating model not only predicts the material quality, but also provides a spatially resolved quality rating. The RAM revealed the negative influence of dislocations and contaminated regions within this application. This technique can be applied to any regression problem and is especially helpful when no human expertise is given.

The second visualization technique allows a human interpretation of the high-dimensional network representation of all data by embedding the data into a 2-D map. The mapping preserves the similarity of the data. The data visualization revealed the continuity and smoothness of the image representation of the model with regard to PL input and V_{oc} output. The map allows a detection of material anomalies or process faults in the case of unexpectedly high prediction errors.

REFERENCES

- [1] T. Trupke, R. Bardos, M. Schubert, and W. Warta, “Photoluminescence imaging of silicon wafers,” *Appl. Phys. Lett.*, vol. 89, no. 4, 2006, Art. no. 044107.
- [2] M. Demant, P. Virtue, A. Kovvali, S. X. Yu, and S. Rein, “Learning quality rating of multicrystalline Si-wafers via convolutional regression networks,” *IEEE, J. Photovolt.*, submitted for publication, doi: [10.1109/JPHOTOV.2019.2906036](https://doi.org/10.1109/JPHOTOV.2019.2906036).
- [3] B. Birkmann, A. Hüsler, A. Seidl, K. Ramspeck, and H. Nagel, “Analysis of multicrystalline wafers originating from corner and edge bricks and forecast of cell properties,” in *Proc. 26th Eur. Photovolt. Sol. Energy Conf. Exhib.*, 2011, pp. 937–940.
- [4] M. Demant *et al.*, “Inline quality rating of multi-crystalline wafers based on photoluminescence images,” *Prog. Photovolt.: Res. Appl.*, vol. 24, no. 12, pp. 1533–1546, 2015.
- [5] T. Trupke and R. A. Bardos, “Wafer imaging and processing method and apparatus,” U.S. Patent 9 546 955, Jan. 17, 2017.
- [6] A. Kovvali, M. Demant, T. Trötschler, J. Haunschild, and S. Rein, “About the relevance of features in as-cut multicrystalline silicon wafers on solar cell performance,” in *Proc. 8th Int. Conf. Crystalline Silicon Photovolt.*, Lausanne, Switzerland, 2018, Art. no. 130011.
- [7] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 818–833.
- [8] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson, “Understanding neural networks through deep visualization,” in *Proc. ICML Deep Learning Workshop*, 2015.
- [9] B. Zhou, A. Khosla, L. Lapedriza, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2921–2929.
- [10] Z. Wang and J. Yang, “Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation,” 2017, arXiv:1703.10757. [Online]. Available: <http://arxiv.org/abs/1703.10757>
- [11] L. van der Maaten and G. E. Hinton, “Visualizing high-dimensional data using t-SNE,” *J. Mach. Learn. Res.*, vol. 9, pp. 2579–2605, 2008.
- [12] J. Donahue *et al.*, “DeCAF: A deep convolutional activation feature for generic visual recognition,” 2013, arXiv:1310.1531.
- [13] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2261–2269.
- [14] C. W. Lan *et al.*, “Grain control in directional solidification of photovoltaic silicon,” *J. Cryst. Growth*, vol. 360, no. 0, pp. 68–75, 2012.
- [15] H. Höffler *et al.*, “Comparison of line-wise PL-imaging and area-wise PL-imaging,” *Energy Procedia*, vol. 124, no. Suppl. C, pp. 66–75, 2017.
- [16] M. A. Green, “The passivated emitter and rear cell (PERC): From conception to mass production,” *Sol. Energy Mater. Sol. Cells*, vol. 143, no. Suppl. C, pp. 190–197, 2015.
- [17] Y. M. Yang *et al.*, “Development of high-performance multicrystalline silicon for photovoltaic industry,” *Prog. Photovolt.: Res. Appl.*, vol. 23, no. 3, pp. 340–351, 2015.
- [18] Y. Jia *et al.*, “Caffe: Convolutional architecture for fast feature embedding,” in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 675–678.

- [19] M. Demant, *Quality Rating of Silicon Wafers—A Pattern Recognition Approach*. Stuttgart, Germany: Fraunhofer-Verlag, 2016.
- [20] M. Glatthar *et al.*, “Evaluating luminescence based voltage images of silicon solar cells,” *J. Appl. Phys.*, vol. 108, 2010, Art. no. 014501.
- [21] M. C. Schubert, H. Habenicht, and W. Warta, “Imaging of metastable defects in silicon,” *IEEE J. Photovolt.*, vol. 1, no. 2, pp. 168–173, Oct. 2011.
- [22] B. Michl *et al.*, “Imaging techniques for quantitative silicon material and solar cell analysis,” *IEEE J. Photovolt.*, vol. 4, no. 6, pp. 1502–1510, Nov. 2014.



Matthias Demant received the Ph.D. degree in engineering sciences from the University of Freiburg, Freiburg, Germany, in 2016.

His thesis was prepared at the Fraunhofer Institute for Solar Energy, Freiburg, as an interdisciplinary work, applying pattern recognition techniques for the characterization of Si wafers and solar cells. As a Postdoctoral Researcher, he joined the Vision Group, International Computer Science Institute, University of California, Berkeley, CA, USA, in 2017. Since 2019, he has been heading a team for “Image Processing and Machine Learning” with the Fraunhofer Institute for Solar Energy

to develop machine learning techniques for material characterization and process control in the field of renewable energy.



Patrick Virtue received the B.S. degree in computer science engineering from the University of Notre Dame, Notre Dame, IN, USA, in 2002. He is working toward the Ph.D. degree in electrical engineering and computer sciences, specializing in deep learning, computer vision, and image processing, with the University of California, Berkeley, CA, USA.

Prior to graduate school, he researched and developed medical image applications as a Software Engineer with GE Healthcare. He has applied his work to a range of applications from medical imaging to

military target detection.



Aditya Kovvali is working toward the Ph.D. degree with Albert Ludwig University, Freiburg, Germany.

He is currently with the Department of Quality Assurance, Characterization and Simulation, Fraunhofer Institute for Solar Energy, Freiburg. His work specializes in data-driven approaches for material assessment in photovoltaics.



Stella X. Yu received the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA, in 2003, where she studied robotics with the Robotics Institute and vision science with the Center for the Neural Basis of Cognition.

She continued her computer vision research as a Postdoctoral Fellow with the University of California, Berkeley, CA, USA, and then studied art and vision as a Clare Boothe Luce Professor with the Boston College. She is currently the Director of Vision Group, International Computer Science Institute, and a Senior Fellow with the Berkeley Institute for Data Science, University of California, Berkeley.

Her research interests include understanding visual perception from multiple perspectives and using computer vision and machine learning to automate and exceed human expertise in practical applications.

Dr. Yu is a recipient of the National Science Foundation CAREER Award.



Stefan Rein studied physics with the University of Freiburg, Freiburg, Germany. He completed the Diploma thesis in the field of defect characterization in semiconductor materials with the Fraunhofer Institute for Solar Energy (Fraunhofer ISE), Freiburg. He received the Ph.D. degree from the University of Konstanz, Konstanz, Germany, in 2004, for his research work on “Lifetime Spectroscopy” done with Fraunhofer ISE.

In 2005, he joined the Industrial Cell Production Group, Fraunhofer ISE, and accompanied design and ramp-up of the Photovoltaic Technology Evaluation Center. He is the Head of the Department “Quality assurance, characterization and simulation,” Fraunhofer ISE. His scientific research interests include defects in crystalline silicon, material and solar cell characterization, development of measurement techniques, and novel approaches for production control.