



Segmentation of photovoltaic module cells in uncalibrated electroluminescence images

Sergiu Deutsch¹ · Claudia Buerhop-Lutz² · Evgenii Sovetkin³ · Ansgar Steland⁴ · Andreas Maier⁵ · Florian Gallwitz⁶ · Christian Riess⁷

Received: 6 August 2019 / Revised: 3 December 2020 / Accepted: 5 March 2021
© The Author(s) 2021

Abstract

High resolution electroluminescence (EL) images captured in the infrared spectrum allow to visually and non-destructively inspect the quality of photovoltaic (PV) modules. Currently, however, such a visual inspection requires trained experts to discern different kinds of defects, which is time-consuming and expensive. Automated segmentation of cells is therefore a key step in automating the visual inspection workflow. In this work, we propose a robust automated segmentation method for extraction of individual solar cells from EL images of PV modules. This enables controlled studies on large amounts of data to understanding the effects of module degradation over time—a process not yet fully understood. The proposed method infers in several steps a high-level solar module representation from low-level ridge edge features. An important step in the algorithm is to formulate the segmentation problem in terms of lens calibration by exploiting the plumbline constraint. We evaluate our method on a dataset of various solar modules types containing a total of 408 solar cells with various defects. Our method robustly solves this task with a median weighted Jaccard index of 94.47% and an F_1 score of 97.62%, both indicating a high sensitivity and a high similarity between automatically segmented and ground truth solar cell masks.

Keywords PV modules · EL imaging · Visual inspection · Lens distortion · Solar cell extraction · Pixelwise classification

1 Introduction

Visual inspection of solar modules using EL imaging allows to easily identify damage inflicted to solar panels either by environmental influences such as hail, during the assembly process, or due to prior material defects or material aging [5,10,65,90,91,93]. The resulting defects can notably decrease the photoelectric conversion efficiency of the modules and thus their energy yield. This can be avoided by continuous inspection of solar modules and maintenance of defective units. For an introduction and review of non-

✉ Sergiu Deutsch
sergiu.deutsch@fau.de

Claudia Buerhop-Lutz
c.buerhop-lutz@fz-juelich.de

Evgenii Sovetkin
e.sovetkin@fz-juelich.de

Ansgar Steland
steland@stochastik.rwth-aachen.de

Andreas Maier
andreas.maier@fau.de

Florian Gallwitz
florian.gallwitz@th-nuernberg.de

Christian Riess
christian.riess@fau.de

¹ Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg, Martensstr. 3, 91058 Erlangen, Germany

² Helmholtz-Institut Erlangen-Nürnberg HI ERN, Forschungszentrum Jülich GmbH, Immerwahrstr. 2, 91058 Erlangen, Germany

³ IEK5-Photovoltaik, Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

⁴ Institute of Statistics, RWTH Aachen University, Wüllnerstr. 3, 52062 Aachen, Germany

⁵ Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg, Martensstr. 3, 91058 Erlangen, Germany

⁶ Faculty of Computer Science, Nuremberg Institute of Technology, Keßlerplatz 12, 90489 Nürnberg, Germany

⁷ IT Security Infrastructures Lab, Friedrich-Alexander University Erlangen-Nürnberg, Martensstr. 3, 91058 Erlangen, Germany

automatic processing tools for EL images, we refer to Mauk [59].

An important step towards an automated visual inspection is the segmentation of individual cells from the solar module. An accurate segmentation allows to extract spatially normalized solar cell images. We already used the proposed method to develop a public dataset of solar cells images [12], which are highly accurate training data for classifiers to predict defects in solar modules [18,60]. In particular, the Convolutional Neural Network (CNN) training is greatly simplified when using spatially normalized samples, because CNNs are generally able to learn representations that are only equivariant to small translations [35, pp. 335–336]. The learned representations, however, are not naturally invariant to other spatial deformations such as rotation and scaling [35,44,52].

The identification of solar cells is additionally required by the international technical specification IEC TS 60904-13 [42, Annex D] for further identification of defects on cell level. Automated segmentation can also ease the development of models that predict the performance of a PV module based on detected or identified failure modes, or by determining the operating voltage of each cell [70]. The data describing the cell characteristics can be fed into an electric equivalent model that allows to estimate or simulate the current-voltage characteristic (I-V) curve [13,46,72] or even the overall power output [47].

The appearance of PV modules in EL images depends on a number of different factors, which makes an automated segmentation challenging. The appearance varies with the type of semiconducting material and with the shape of individual solar cell wafers. Also, cell cracks and other defects can introduce distracting streaks. A solar cell completely disconnected from the electrical circuit will also appear much darker than a functional cell. Additionally, solar modules vary in the number of solar cells and their layout, and solar cells themselves are oftentimes subdivided by busbars into multiple segments of different sizes. Therefore, it is desirable for a fully automated segmentation to infer both the arrangement of solar cells within the PV module and their subdivision from EL images alone, in a way that is robust to various disturbances. In particular, this may ease the inspection of heterogeneous batches of PV modules.

In this work, we assume that EL images are captured in a manufacturing setting or under comparable conditions in a test laboratory where field-aged modules are analyzed either regularly or after hazards like hailstorms. Such laboratories oftentimes require agile work processes where the equipment is frequently remounted. In these scenarios, the EL irradiation of the solar module predominates the background irradiation, and the solar modules are captured facing the EL camera without major perspective distortion. Thus, the geometric distortions that are corrected by the proposed method are radial lens distortion, in-plane rotation, and minor per-

spective distortions. This distinguishes the manufacturing setting from acquisitions in the field, where PV modules may be occluded by cables and parts of the rack, and the perspective may be strong enough to require careful correction. However, perspective distortion also makes it more difficult to identify defective areas (e.g., microcracks) due to the foreshortening effect [4]. Therefore, capturing EL images from an extreme perspective is generally not advisable. Specifically for manufacturing environments, however, the proposed method yields a robust, highly accurate, and completely automatic segmentation of solar modules into solar cells from high resolution EL images of PV modules.

Independently of the setting, our goal is to allow for some flexibility for the user to freely position the camera or use zoom lenses without the need to recalibrate the camera.

With this goal in mind, a particular characteristic of the proposed segmentation pipeline is that it does not require an external calibration pattern. During the detection of the grid that identifies individual solar cells, the busbars and the inter solar cell borders are directly used to estimate lens distortion. Avoiding the use of a separate calibration pattern also avoids the risk of an operator error during the calibration, e.g., due to inexperienced personnel.

A robust and fully automatic PV module segmentation can help understanding the influence of module degradation on module efficiency and power generation. Specifically, this allows to continuously and automatically monitor the degradation process, for instance, by observing the differences in a series of solar cell images captured over a certain period of time. The segmentation also allows to automatically create training data for learning-based algorithms for defect classification and failure prediction.

1.1 Contributions

To the best of our knowledge, the proposed segmentation pipeline is the first work to enable a fully automatic extraction of solar cells from uncalibrated EL images of solar modules (cf., Fig. 1b). Within the pipeline, we seek to obtain the exact segmentation mask of each solar cell through estimation of nonlinear and linear transformations that warp the EL image into a canonical view. To this end, our contributions are three-fold:

1. Joint camera lens distortion estimation and PV module grid detection for precise solar cell region identification.
2. A robust initialization scheme for the employed lens distortion model.
3. A highly accurate pixelwise classification into active solar cell area on monocrystalline and polycrystalline PV modules robust to various typical defects in solar modules.

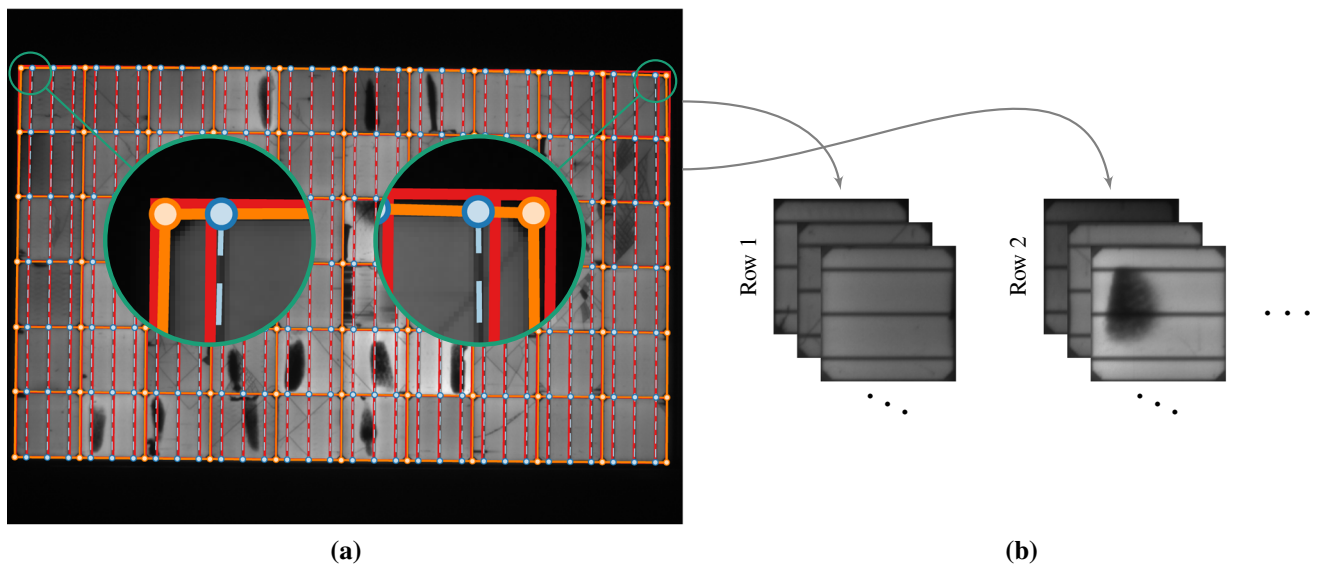


Fig. 1 (a) An EL image of a PV module overlaid by a rectangular grid (—) and parabolic curve grid (—○—) determined using our approach. The intersections of the rectangular grid were registered to curve grid intersections to accurately align both grids. Notice how the rectangular grid is still not able to capture the curved surface of the solar module induced by the (weak) lens distortion that increases especially towards the image border. Using the curve grid, we estimate the lens distortion, rectify the image and finally extract the individual cells using the estimated module topology (b). The segmented solar cells can be used for further analysis, such as automatic defect classification or failure prediction in PV modules. The solar cells are approximately $15.6 \text{ cm} \times 15.6 \text{ cm}$ with a standard 60 cell PV module with overall dimensions of $1 \text{ m} \times 1.65 \text{ m}$

Moreover, our method operates on arbitrary (unseen) module layouts without prior knowledge on the layout.

1.2 Outline

The remainder of this work is organized as follows. Section 2 discusses the related work. In Sect. 3, the individual stages of the segmentation pipeline are presented. In Sect. 4, we evaluate the presented segmentation approach on a number of different PV modules with respect to the segmentation accuracy. Finally, the conclusions are given in Sect. 5.

2 Related work

The segmentation of PV modules into individual solar cells is related to the detection of calibration patterns, such as checkerboard patterns commonly used for calibrating intrinsic camera and lens parameters [29,36,41,69,79]. However, the appearance of calibration patterns is typically perfectly known, whereas detection of solar cells is encumbered by various defects that are a priori unknown. Additionally, the number of solar cells in a PV module and their layout can vary. We also note that existing lens models generally assume wide angle lenses. However, their application to standard lenses is to our knowledge not widely studied.

To estimate the parameters of a lens distortion model, the plumline constraint is typically employed [11]. The

constraint exploits the fact that the projection of straight lines under radial and tangential distortion will not be truly straight. For example, under radial distortion, straight lines are images as curves. For typical visual inspection tasks, a single image is sufficient to estimate the lens distortion parameters [2,16,17,20,25,78]. This can be achieved by decoupling the intrinsic parameters of the camera from the parameters of the lens distortion model [20].

Novel methodologies employ CNNs for various segmentation tasks. Existing CNN-based segmentation tasks can be categorized into (1) object detection, (2) semantic segmentation, and (3) instance-aware segmentation. One of the first CNN object detection architectures is Regions with CNN features (R-CNN) [32] to learn features that are subsequently classified using a class-specific linear Support Vector Machine (SVM) to generate region proposals. R-CNN learns to simultaneously classify object proposals and refine their spatial locations. The predicted regions, however, provide only a coarse estimation of object's location in terms of bounding boxes. Girshick [31] proposed Fast Region-based Convolutional Neural Network (Fast R-CNN) by accelerating training and testing times while also increasing the detection accuracy. Ren et al. [75] introduced Region Proposal Network (RPN) that shares full-image convolutional features with the detection network enabling nearly cost-free region proposals. RPN is combined with Fast R-CNN into a single network that simultaneously predicts object bounds and estimates the probability of an object for each

proposal. For semantic segmentation, Long et al. [56] introduced Fully Convolutional Networks (FCNs) allowing for pixelwise inference. The FCN is learned end-to-end and pixels-to-pixels requiring appropriately labeled training data. Particularly, in medical imaging the U-Net network architecture by Ronneberger et al. [77] has been successfully applied for various segmentation tasks. In instance segmentation, Li et al. [51] combined segment proposal and object detection for Fully Convolutional Instance Segmentation (FCIS) where the general idea is to predict the locations in a fully convolutional network. He et al. [39] proposed a Mask R-CNN which extends Faster R-CNN.

The work by Mehta et al. [62] introduces a CNN for the prediction of power loss. Their system additionally localizes and classifies the type of soiling. Their work is based on RGB images of whole PV modules and addresses the additional geometric challenges of acquisitions in the field. In contrast, this work operates on EL images of individual cells of a PV module, and in particular focuses on their precise segmentation in a manufacturing setting.

The main limitation of learning-based approaches is the requirement of a considerable number of appropriately labeled images for training. However, pixelwise labeling is time-consuming, and in absence of data not possible at all. Also, such learning-based approaches require training data that is statistically representative for the test data, which oftentimes requires to re-train a model on data with different properties. In contrast, the proposed approach can be readily deployed to robustly segment EL images of PV modules without notable requirements of labeled training data.

The closest work related to the proposed method was presented by Sovetkin and Steland [86]. This method proposes a robust PV module grid alignment for the application on field EL images, where radial and perspective distortion, motion blur, and disturbing background may be present. The method uses an external checkerboard calibration for radial distortion correction, and prior knowledge on the solar cell topology in terms of the relative distances of the grid lines separating the busbars and cell segments. In contrast, EL images taken under manufacturing conditions may be cropped or rotated, and the camera is not always pre-calibrated. Hence, the proposed method performs an automated on-line calibration for every EL image. This is particularly useful for EL images of PV modules from various sources, for which the camera parameters may not be available, or when zoom lenses are used. Additionally, the proposed method performs a pixelwise classification of pixels belonging to the active cell area and therefore is able to provide masks tailored to a specific module type. Such masks allow to exclude unwanted background information and to simplify further processing.

In this work, we unify lens distortion estimation and grid detection by building upon ideas of Devernay and Faugeras [20]. However, instead of using independent line segments to

estimate lens distortion parameters, we constrain the problem using domain knowledge by operating on a coherent grid. This joint methodology allows to correct errors through feedback from the optimization loop used for estimating lens model parameters. The proposed approach conceptually differs from Sovetkin and Steland [86], where both steps are decoupled and an external calibration is required.

3 Methodology

The proposed framework uses a bottom-up pipeline to gradually infer a high-level representation of a solar module and its cells from low-level ridge edge features in an EL image. Cell boundaries and busbars are represented as parabolic curves to robustly handle radial lens distortion which causes straight lines to appear curved in the image. Once we estimated the lens distortion parameters, the parabolas are rectified to obtain a planar cell grid. This rectified representation is used to segment the solar cells.

3.1 Overview

The general framework for segmenting the solar cells in EL images of PV modules is illustrated in Fig. 2 and consists of the following steps. First, we locate the busbars and the inter solar cell borders by extracting the ridge edges. The ridge edges are extracted at subpixel accuracy and approximated by a set of smooth curves defined as second-degree polynomials. The parametric representation is used to construct an initial grid of perpendicularly arranged curves that identify the PV module. Using this curve grid, we estimate the initial lens distortion parameters and hypothesize the optimal set of curves by further excluding outliers in a RANdom SAMple Consensus (RANSAC) scheme. Then we refine the lens distortion parameters that we eventually use to rectify the EL image. From the final set of curves we infer the PV module configuration and finally extract the size, perspective, and orientation of solar cells.

3.2 Preprocessing

First, the contrast of an EL image is enhanced to account for possible underexposure. Then, low-level edge processing is applied to attenuate structural variations that might stem from cracks or silicon wafer texture, with the goal of preserving larger lines and curves.

3.2.1 Contrast enhancement

Here, we follow the approach by Franken et al. [28]. A copy I_{bg} of the input EL image I is blurred with a Gaussian kernel, and a morphological closing with a disk-shaped

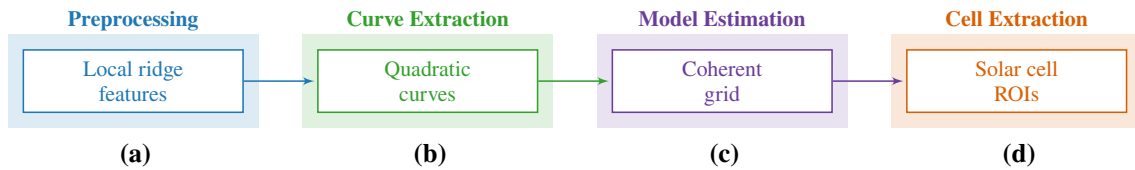


Fig. 2 The proposed PV module segmentation pipeline consists of four stages. In the preprocessing stage (a), local ridge features are extracted. In the curve extraction stage (b), candidate parabolic curves are determined from ridges. In the model estimation stage (c), a coherent grid and the lens distortion are jointly estimated. In the cell extraction stage (d) the cell topology is determined and the cells are extracted

structure element is applied. Dividing each pixel of I by I_{bg} attenuates unwanted background noise while emphasizing high contrast regions. Then, histogram equalization [34, pp. 134sq.] is applied to increase its overall contrast. Figure 5b shows the resulting image I .

3.2.2 Gaussian scale-space ridgeness

The high-level grid structure of a PV module is defined by inter-cell borders and busbars, which correspond to ridges in the image. Ridge edges can be determined from second-order partial derivatives summarized by a Hessian. To robustly extract line and curve ridges, we compute the second-order derivative of the image at multiple scales [54,55]. The responses are computed in a Gaussian pyramid constructed from an input EL image [53]. This results in several layers of the pyramid at varying resolutions commonly referred to as octaves. The eigendecomposition of the Hessian computed afterwards provides information about line-like structures.

More in detail, let $\mathbf{u} := (u, v)^T$ denote discrete pixel coordinates, $O \in \mathbb{N}$ the number of octaves in the pyramid, and $P \in \mathbb{N}$ the number of sublevels in each octave. At the finest resolution, we set σ to the golden ratio $\sigma = 1 + \sqrt{5}/2 \approx 1.6$. At each octave $o \in \{0, \dots, O-1\}$ and sublevel $\ell \in \{0, \dots, P-1\}$, we compute the Hessian by convolving the image with the derivatives of the Gaussian kernel. To obtain the eigenvalues, the symmetric Hessian is diagonalized by annihilating the off-diagonal elements using the Jacobi method which iteratively applies Givens rotations to the matrix [33]. This way, its eigenvalues and the corresponding eigenvectors can be simultaneously extracted in a numerically stable manner. Let $\mathbf{H} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T$ denote the eigendecomposition of the Hessian \mathbf{H} , where $\mathbf{\Lambda} := \text{diag}(\lambda_1, \lambda_2) \in \mathbb{R}^{2 \times 2}$ is a diagonal matrix of eigenvalues $\lambda_1 > \lambda_2$ and $\mathbf{V} := (\mathbf{v}_1, \mathbf{v}_2)$ are the associated eigenvectors. Under a Gaussian assumption, the leading eigenvector dominates the likelihood if the associated leading eigenvalue is spiked. In this sense, the local ridgeness describes the likelihood of a line segment in the image at position \mathbf{u} , and the orientation of the associated eigenvector specifies the complementary angle $\beta(\mathbf{u})$ of the most likely line segment orientation at position \mathbf{u} . The local ridgeness $R(\mathbf{u})$ is obtained as the maximum positive eigenvalue $\lambda_1(\mathbf{u})$ across all octaves and sublevels. Both the ridgeness $R(\mathbf{u})$ and

the angle $\beta(\mathbf{u})$ provide initial cues for ridge edges in the EL image (see Fig. 5c).

3.2.3 Contextual enhancement via tensor voting

Ridgeness can be very noisy (cf., Fig. 5c). To discern noise and high curvatures from actual line and curve features, $R(\mathbf{u})$ is contextually enhanced using tensor voting [61].

Tensor voting uses a stick tensor voting field to model the likelihood that a feature in the neighborhood belongs to the same curve as the feature in the origin of the voting field [27]. The parameter $\varsigma > 0$ controls the proximity of the voting field, and ν determines the angular specificity that we set to $\nu = 2$ in our experiments.

Following Franken et al. [27], stickness $\tilde{R}(\mathbf{u}) = \tilde{\lambda}_1 - \tilde{\lambda}_2$ is computed as the difference between the two eigenvalues $\tilde{\lambda}_1, \tilde{\lambda}_2$ of the tensor field, where $\tilde{\lambda}_1 > \tilde{\lambda}_2$. $\tilde{\beta}(\mathbf{u}) = \angle \tilde{\mathbf{e}}_1$ is the angle of the eigenvector $\tilde{\mathbf{e}}_1 \in \mathbb{R}^2$ associated with the largest eigenvalue $\tilde{\lambda}_1$, analogously to $\beta(\mathbf{u})$.

We iterate tensor voting two times, since one pass is not always sufficient [28]. Unlike Franken et al., however, we do not thin out the stickness immediately after the first pass to avoid too many disconnected edges. Given the high resolution of the EL images in our dataset of approximately 2500×2000 pixels, we use a fairly large proximity of $\varsigma_1 = 15$ in the first tensor voting step, and $\varsigma_2 = 10$ in the second.

Figure 5d shows a typical stickness $\tilde{R}(\mathbf{u})$ output. The stickness along the orientation $\tilde{\beta}(\mathbf{u})$ is used to extract curves at subpixel accuracy in the next step of the pipeline.

3.3 Curve extraction

We seek to obtain a coherent grid which we define in terms of second-degree curves. These curves are traced along the previously extracted ridges by grouping centerline points by their curvature. We then fit second-degree polynomials to these points, which yields a compact high-level curve representation while simultaneously allowing to discard point outliers.

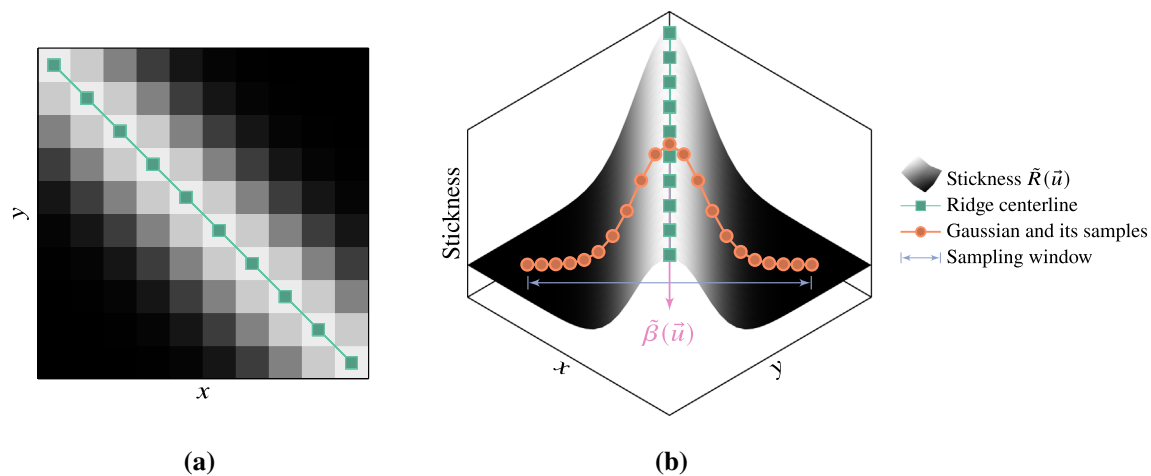


Fig. 3 Extraction of ridge edges from stickness at subpixel accuracy. (a) shows a stickness patch with its initial centerline (—■—) at discrete coordinates obtained by skeletonization. The refined ridge centerline at subpixel accuracy is estimated by fitting a Gaussian function (—●—) to the cross-section profile of the ridge edge in (b) to equidistantly sampled stickness values within a predefined sampling window (—↔—)

3.3.1 Extraction of ridges at subpixel accuracy

To ensure a high estimation accuracy of lens distortion parameters, we extract ridge edges at subpixel accuracy. This also makes the segmentation more resilient in out-of-focus scenarios, where images may appear blurry and the ridge edges more difficult to identify due to their smoother appearance. Blurry images can be caused by slight camera vibrations during the long exposure time of several seconds that is required for imaging. Additionally, focusing in a dark room can be challenging, hence blur cannot be always avoided. Nevertheless, it is beneficial to be able to operate also on blurry images, as they can still be useful for defect classification and power yield estimation in cell areas that do not irradiate.

To this end, we perform non-maximum suppression by Otsu's global thresholding [67] on the stickness $\tilde{R}(\vec{u})$ followed by skeletonization [80]. Afterwards, we collect the points that represent the centerline of the ridges through edge linking [48]. The discrete coordinates can then be refined by setting the centerline to the mean of a Gaussian function fitted to the edge profile [23] using the Gauss-Newton (GN) optimization algorithm [66]. The 1-dimensional window of the Gaussian is empirically set to 21 pixels, with four sample points per pixel that are computed via bilinear interpolation. The GN algorithm is initialized with the sample mean and standard deviation in the window, and multiplicatively scaled to the stickness magnitude at the mean. The mean of the fitted Gaussian is then reprojected along the edge profile oriented at $\tilde{\beta}(\vec{u})$ to obtain the edge subpixel position. Figure 3 visualizes these steps.

A nonparametric alternative to fitting a Gaussian to the ridge edge profile constitutes fitting a parabola instead [19]. Such an approach is very efficient since it involves a closed-

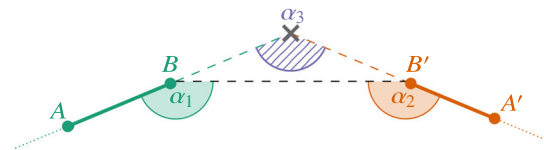


Fig. 4 When considering combining two adjacent curve segments, one with the end line segment \vec{AB} and the other with the start line segment $\vec{B'A'}$, we evaluate the angles α_1 , α_2 , and α_3 and ensure they are below the predefined threshold ϑ with $\alpha_1, \alpha_2 \geq \alpha_3 \geq \pi - \vartheta$. This way, the combined curve segments are ensured to have a consistent curvature

form solution. On the downside, however, the method suffers from oscillatory artifacts which require additional treatment [30].

3.3.2 Connecting larger curve segments

A limitation of the edge linking method is that it does not prioritize curve pairs with similar orientation. To address this, we first reduce the set of points that constitute a curve to a sparse representation using the nonparametric variant of the Ramer-Douglas-Peucker algorithm [21, 73] introduced by Prasad et al. [71]. Afterwards, edges are disconnected if the angle between the corresponding line segments is nonzero. In a second pass, two line segments are joined if they are nearby, of approximately the same length, and pointing into the same direction within an angle range $\vartheta = 5^\circ$. Figure 4 illustrates the way two curve segments are combined.

In the final step, the resulting n_i points of the i -th curve of a line segment form a matrix $\hat{Q}^{(i)} \in \mathbb{R}^{2 \times n_i}$. For brevity, we denote the j -th column of $\hat{Q}^{(i)}$ by $\hat{q}_j \in \mathbb{R}^2$. $\hat{Q}^{(i)}$ is used to find the parametric curve representation.

3.3.3 Parametric curve representation

Projected lines are represented as second-degree polynomials to model radial distortion. The curve parameters are computed via linear regression on the curve points. More specifically, let

$$f(x) = a_2x^2 + a_1x + a_0 \quad (1)$$

denote a second-degree polynomial in horizontal or vertical direction. The curve is fitted to line segment points $\hat{\mathbf{q}}_j \in \{(x_j, y_j)^\top \mid j = 1, \dots, n_i\} \subseteq \hat{\mathbf{Q}}^{(i)}$ of the i -th curve $\hat{\mathbf{Q}}^{(i)}$ by minimizing the Mean Squared Error (MSE)

$$\text{MSE}(f) = \frac{1}{n_i} \sum_{j=1}^{n_i} (f(x_j) - y_j)^2 \quad (2)$$

using RANSAC iterations [24]. In one iteration, we randomly sample three points to fit Eq. (1), and then determine which of the remaining points support this curve model via MSE. Outlier points are discarded if the squared difference between the point and the parabolic curve value at its position exceeds $\rho = 1.5$. To keep the computational time low, RANSAC is limited to 100 iterations, and stopped early once sufficiently many inliers at a 99 % confidence level are found [38, ch. 4.7]. After discarding the outliers, each curve is refitted to supporting candidate points using linear least squares [33]. To ensure a numerically stable and statistically robust fit, the 2-D coordinates are additionally normalized [37].

3.4 Curve grid model estimation

The individual curves are used to jointly form a grid, which allows to further discard outliers, and to estimate lens distortion. To estimate the lens distortion, we employ the plumblines constraint [11]. The constraint models the assumption that curves in the image correspond to straight lines in real world. In this way, it becomes possible to estimate distortion efficiently from a single image, which allows to use this approach also post hoc on cropped, zoomed or similarly processed images.

3.4.1 Representation of lens distortion

Analogously to Devernay and Faugeras [20], we represent the radial lens distortion by a function $L: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ that maps the distance of a pixel from the distortion center to a distortion factor. This factor can be used to radially displace each normalized image coordinate $\tilde{\mathbf{x}}$.

Image coordinates are normalized by scaling down coordinates $\mathbf{x} := (x, y)^\top$ horizontally by the distortion aspect ratio s_x (corresponding to image aspect ratio decoupled from

the projection on the image plane) followed by shifting the center of distortion $\mathbf{c} := (c_x, c_y)^\top$ to the origin and normalizing the resulting 2-D point to the unit range using the dimensions $M \times N$ of the image of width M and height N . Homogeneous coordinates allow to express the normalization conveniently using a matrix product. By defining the upper-triangular matrix

$$\mathbf{K} = \begin{bmatrix} s_x M & 0 & c_x \\ 0 & N & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

the normalizing mapping $\mathbf{n}: \Omega \rightarrow [-1, 1]^2$ is

$$\mathbf{n}(\mathbf{x}) = \boldsymbol{\pi} \left(\mathbf{K}^{-1} \boldsymbol{\pi}^{-1}(\mathbf{x}) \right), \quad (4)$$

where $\boldsymbol{\pi}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ projects homogeneous to inhomogeneous coordinates,

$$\boldsymbol{\pi}: (x, y, z)^\top \mapsto \frac{1}{z}(x, y)^\top, \quad \text{for } z \neq 0 \quad (5)$$

and the inverse operation $\boldsymbol{\pi}^{-1}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ backprojects inhomogeneous to homogeneous coordinates:

$$\boldsymbol{\pi}^{-1}: (x, y)^\top \mapsto (x, y, 1)^\top. \quad (6)$$

Note that the inverse mapping \mathbf{n}^{-1} converts normalized image coordinates to image plane coordinates.

3.4.2 The field-of-view lens distortion model

To describe the radial lens distortion, we use the first-order Field-of-View (FOV) lens model by Devernay and Faugeras that has a single distortion parameter ω . While images can also suffer from tangential distortion, this type of distortion is often negligible [92]. The sole parameter $0 < \omega \leq \pi$ denotes the opening angle of the lens. The corresponding radial displacement function L is defined in terms of the distortion radius $r \geq 0$ as

$$L(r) = \frac{1}{\omega} \arctan \left(2r \tan \frac{\omega}{2} \right), \quad \text{for } \omega \neq 0. \quad (7)$$

One advantage of the model is that its inversion has a closed-form solution with respect to the distortion radius r .

Similar to Devernay and Faugeras, we decouple the distortion from the projection onto the image plane, avoiding the need to calibrate for intrinsic camera parameters. Instead, the distortion parameter ω is combined with the distortion center $\mathbf{c} \in \Omega$ and distortion aspect ratio s_x which are collected in a vector $\boldsymbol{\theta} := (\mathbf{c}, s_x, \omega)$.

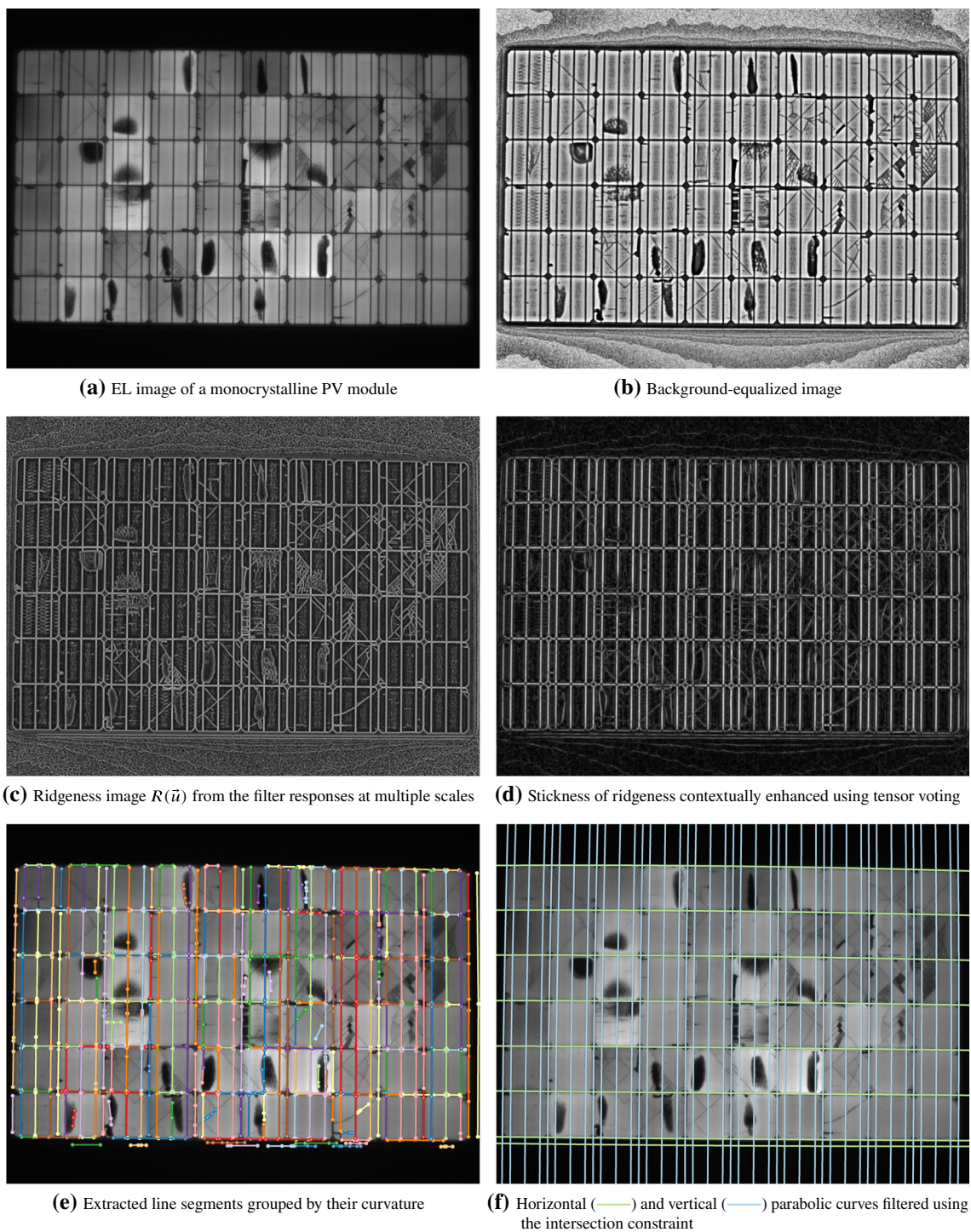


Fig. 5 Visualization of the preprocessing, curve extraction, and model estimation stages for the PV module from Fig. 1

Normalized undistorted image coordinates $\tilde{\mathbf{x}}_u = \delta^{-1}(\tilde{\mathbf{x}}_d)$ can be directly computed from distorted coordinates $\tilde{\mathbf{x}}_d$ as

$$\delta^{-1}(\tilde{\mathbf{x}}_d) = \frac{L^{-1}(r_d)}{r_d} \tilde{\mathbf{x}}_d, \quad \text{for } r_d \neq 0 \quad (8)$$

where $r_d = \|\tilde{\mathbf{x}}_d\|_2$ is the distance of $\tilde{\mathbf{x}}_d$ from the origin. $L^{-1}(r)$ is the inverse of the lens distortion function in Eq. (7), namely

$$L^{-1}(r) = \frac{\tan r \omega}{2 \tan \frac{\omega}{2}}, \quad \text{for } \omega \neq 0. \quad (9)$$

The function that undistorts a point $\mathbf{x} \in \Omega$ is thus

$$\mathbf{u}(\mathbf{x}) = \mathbf{n}^{-1}(\delta^{-1}(\mathbf{n}(\mathbf{x}))) . \quad (10)$$

Note that Eq. (8) exhibits a singularity at $r_d \approx 0$ for points close to the distortion center. By inspecting the function's limits, one obtains

$$\lim_{r_d \rightarrow 0^+} \delta^{-1}(\tilde{\mathbf{x}}_d) = \frac{\omega}{2 \tan \frac{\omega}{2}} \tilde{\mathbf{x}}_d . \quad (11)$$

Analogously, Eq. (9) is singular at $\omega = 0$ but approaches $\lim_{r \rightarrow 0^+} L^{-1}(r) = r$ at the limit. In this case, Eq. (8) is an identity transformation which does not radially displace points.

3.4.3 Estimation of initial lens distortion model parameters

Lens distortion is specified by the distortion coefficient ω , the distortion aspect ratio s_x , and the distortion center \mathbf{c} . Naive solution leads to a non-convex objective function with several local minima. Therefore, we first seek an initial set of parameters close to the optimum, and then proceed using a convex optimization to refine the parameters.

We propose the following initialization scheme for the individual parameters of the FOV lens model.

Distortion Aspect Ratio and Center We initialize the distortion aspect ratio to $s_x = 1$, and the distortion center to the intersection of two perpendicular curves with smallest coefficients in the highest order polynomial term. Such curves can be assumed to have the smallest curvature and are thus located near the distortion center.

To find the intersection of two perpendicular curves, we denote the coefficients of a horizontal curve by a_2, a_1, a_0 , and the coefficients of a vertical curve by b_2, b_1, b_0 . The position x of a curve intersection is then the solution to

$$a_2^2 b_2 x^4 + 2a_1 a_2 b_2 x^3 + x^2 (2a_0 a_2 b_2 + a_1^2 b_2 + a_2 b_1) + x \cdot (2a_0 a_1 b_2 + a_1 b_1 - 1) + a_0^2 b_2 + a_0 b_1 + b_0 = 0 . \quad (12)$$

The real roots of the quartic (12) can be found with the Jenkins-Traub RPOLY algorithm [45] or a specialized quartic solver [26]. The corresponding values $f(x)$ are determined by inserting the roots back into Eq. (1).

Distortion Coefficient Estimation of the distortion coefficient ω from a set of distorted image points is not straightforward because the distortion function $L(r)$ is nonlinear. One way to overcome this problem is to linearize $L(r)$ with Taylor polynomials, and to estimate ω with linear least squares.

To this end, we define the distortion factor

$$k := \frac{L(r)}{r}, \quad \text{for } k \in \mathbb{R}_{>0} \quad (13)$$

which maps undistorted image points $\{\mathbf{p}_j\}_{j=1}^n$ lying on the straight lines to distorted image points $\{\mathbf{q}_j\}_{j=1}^n$ lying on the parabolic curves. Both point sets are then related by

$$\mathbf{p}k = \mathbf{q} . \quad (14)$$

The distorted points \mathbf{q}_j are straightforward to extract by evaluating the second-degree polynomial of the parabolic curves. To determine \mathbf{p}_j , we define a line with the first and the last point in \mathbf{q}_j , and select points from this line. Collecting these points in the vectors $\mathbf{p} \in \mathbb{R}^{2n}$ and $\mathbf{q} \in \mathbb{R}^{2n}$ yields an overdetermined system of $2n$ linear equations in one unknown. \hat{k} is then estimated via linear least squares as

$$\hat{k} = \underset{k}{\operatorname{argmin}} \|\mathbf{q} - \mathbf{p}k\|_2^2, \quad (15)$$

where the solution is found via the normal equations [33] as

$$\hat{k} := \frac{\mathbf{p}^\top \mathbf{q}}{\mathbf{p}^\top \mathbf{p}} . \quad (16)$$

The points $\mathbf{q}_j, \mathbf{p}_j$ refer to the columns of the two matrices $\mathbf{Q}^{(i)}, \mathbf{P}^{(i)} \in \mathbb{R}^{2 \times n_i}$, respectively, where n_i again denotes the number of points, which are used in the following step of the pipeline.

To determine ω from the relation $k = \frac{L(r)}{r}$, $L(r)$ is expanded around $\omega_0 = 0$ using Taylor series. More specifically, we use a second-order Taylor expansion to approximate

$$\arctan(x) = x + \mathcal{O}(x^2), \quad (17)$$

and a sixth-order Taylor expansion to approximate

$$\tan(y) = y + \frac{y^3}{3} + \frac{2y^5}{15} + \mathcal{O}(y^6) . \quad (18)$$

Let $L(r) = \frac{1}{\omega} \arctan(x)$ with $x = 2r \tan(y)$, and $y = \frac{\omega}{2}$. We substitute the Taylor polynomials from Eqs. (17) and

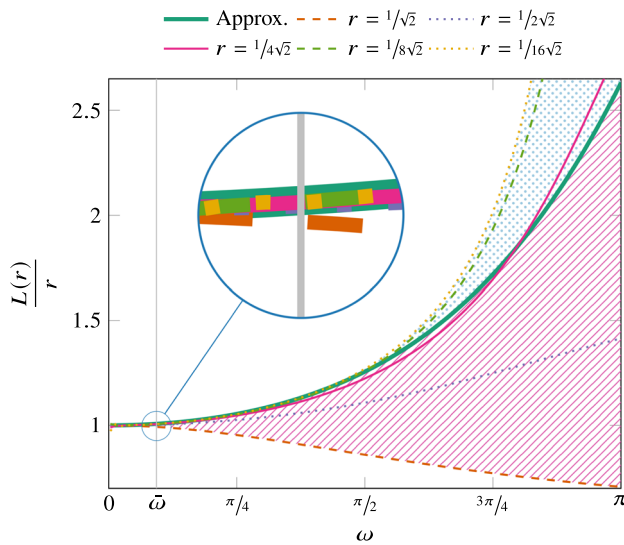


Fig. 6 Approximation of the distortion coefficient ω using Eq. (19) (—) compared to the exact solution with respect to varying radii r . For large radii outside the range of normalized coordinates (i.e., the radius of the half-unit circle $r > 1/\sqrt{2}$), the estimate is not accurate. This implies that the ideal sampled points must be both at some distance from the image border and also from the distortion center. As a side note, the estimation error becomes unacceptable for wide lenses where $\omega > \pi/4$. However, the EL images in this work ($\bar{\omega}$) are well below this threshold

(18), and x, y into Eq. (13) to obtain a biquadratic polynomial $Q(\omega)$ independent of r :

$$\frac{L(r)}{r} \approx 1 + \underbrace{\frac{1}{12}\omega^2 + \frac{1}{120}\omega^4}_{=: Q(\omega)}. \quad (19)$$

By equating the right-hand side of Eq. (19) to k

$$Q(\omega) = k \quad (20)$$

we can estimate ω from four roots of the resulting polynomial $Q(\omega)$. These roots can be found by substituting $z = \omega^2$ into Eq. (19), solving the quadratic equation with respect to z , and substituting back to obtain ω . This eventually results in the four solutions $\pm\sqrt{z_{1,2}}$. The solution exists only if $k \geq 1$, as complex solutions are not meaningful, and thus corresponds to the largest positive real root.

We evaluated the accuracy of the approximation (19) with the results shown in Fig. 6. For large radii, the approximation significantly deviates from the exact solution. Consequently, this means that the selected points for the estimation must ideally be well distributed across the image. Otherwise, the lens distortion parameter will be underestimated. In practice, however, this constraint does not pose an issue due to the spatial distribution of the solar cells across the captured EL image.

3.4.4 Minimization criterion for the refinement of lens distortion parameters

The Levenberg-Marquardt algorithm [50,57] is used to refine the estimated lens distortion parameters θ . The objective function is

$$\theta^* := \underset{\theta}{\operatorname{argmin}} \frac{1}{2} \sum_{i=1}^n \chi^2(P^{(i)}, \theta). \quad (21)$$

$P^{(i)} \in \mathbb{R}^{2 \times m}$ is a matrix of m 2-D points of the i -th curve. The distortion error χ^2 quantifies the deviation of the points from the corresponding ideal straight line [20]. The undistorted image coordinates $\mathbf{p}_j := (x_j, y_j)^\top \in \Omega$ are computed as $\mathbf{p}_j = \mathbf{u}(\mathbf{q}_j)$ by applying the inverse lens distortion given in Eq. (10) to the points \mathbf{q}_j of the i -th curve $Q^{(i)}$. In a similar manner, the obtained points \mathbf{p}_j form the columns of $P^{(i)} \in \mathbb{R}^{2 \times n_i}$.

Following Devernay and Faugeras, we iteratively optimize the set of lens parameters θ . In every step t , we refine these parameters and then compute the overall error $\epsilon_t := \sum_{i=1}^n \chi^2(P^{(i)}, \theta)$ over all curve points. Afterwards, we undistort the curve points and continue the optimization until the relative change in error $\epsilon := (\epsilon_{t-1} - \epsilon_t)/\epsilon_t$ falls below the threshold $\epsilon = 10^{-6}$.

Minimizing the objective function (21) for all parameters simultaneously may cause the optimizer to be trapped in a local minimum. Hence, following Devernay and Faugeras [20], we optimize the parameters $\theta = (\omega, s_x, \mathbf{c})$ in subsets starting with ω only. Afterwards, we additionally optimize the distortion center \mathbf{c} . Finally, the parameters θ are jointly optimized.

3.4.5 Obtaining a consistent parabolic curve grid model

The layout of the curves is constrained to a grid in order to eliminate outlier curves. Ideally, each horizontally oriented parabola should intersect each vertically oriented parabola exactly once. This intersection can be found using Eq. (12). Also, every parabolic curve should not intersect other parabolic curves of same orientation within the image plane. This set of rules eliminates most of the outliers.

Robust Outlier Elimination Locally Optimized RANDOM Sample Consensus (LO-RANSAC) [15] is used to remove outlier curves. In every LO-RANSAC iteration, the grid constraints are imposed by randomly selecting two horizontal and two vertical curves to build a minimal grid model. Inliers are all curves that (1) exactly once intersect the model grid lines of perpendicular orientation, (2) not intersect the model grid lines of parallel orientation, and (3) whose MSE of the reprojected undistorted points is not larger than one pixel.

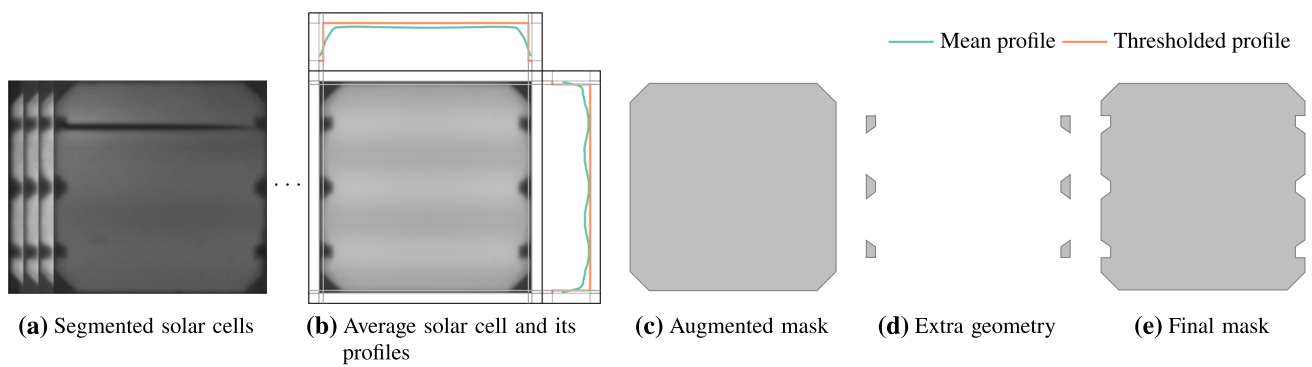


Fig. 8 Intermediate steps of the solar mask estimation process

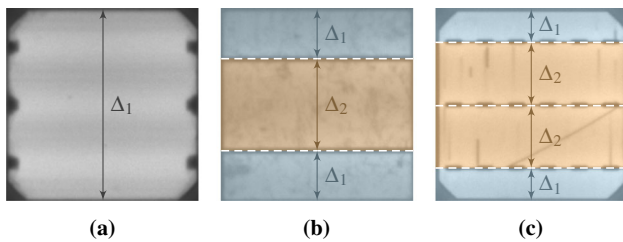


Fig. 7 Estimation of solar module topology requires determining the number of subdivisions (i.e., rectangular segments) in a solar cell. Common configurations include no subdivisions at all (i.e., one segment) (a), three segments (b) and four segments (c). Notice how the arrangement of rectangular segments is symmetric and segment sizes increase monotonically towards the center, i.e., $\Delta_1 < \dots < \Delta_n$. In particular, shape symmetry can be observed not only along the vertical axis of the solar cell but also along the horizontal one as well

Remaining Curve Outliers Halos around the solar modules and holding mounts (such as in Fig. 5) can generate additional curves outside of the cells. We apply Otsu's thresholding [67] on the contrast-normalized image and discard outer curves that generate additional grid rows or columns with an average intensity in the enclosed region below the automatically determined threshold.

3.5 Estimation of the solar module topology

A topology constraint on the solar cell can be employed to eliminate remaining non-cell curves in the background of the PV module, and the number and layout of solar cells can be subsequently estimated. However, outliers prevent a direct estimation of the number of solar cell rows and columns in a PV module. Additionally, the number and orientation of segments dividing each solar cell are generally unknown. Given the aspect ratio of solar cells in the imaged PV module, the topology can be inferred from the distribution of parabolic curves. For instance, in PV modules with equally long horizontal and vertical cell boundary lines, the solar cells have a square (i.e., 1 : 1) aspect ratio.

The number of curves crossing each square image area of solar cell is constant. Clustering the distances between the curves allows to deduce the number of subdivisions within solar cells.

3.5.1 Estimation of the solar cell subdivisions and the number of rows and columns

The solar cells and their layout are inferred from the statistics of the line segment lengths in horizontal and vertical direction. We collect these lengths separately for each dimension and cluster them. DBSCAN clustering [22] is used to simultaneously estimate cluster membership and the number of clusters. Despite the presence of outlier curves, clusters are representative of the distribution of segment dimensions within a cell. For example, if a solar cell consists of three vertically arranged segments (as in Fig. 7b) with heights of 20 : 60 : 20 pixels, the two largest clusters will have the medians 60 and 20. With the assumption that the segment arrangement is typically symmetric, the number of segments is estimated as the number of clusters times two minus one. If clustering yields a single cluster, we assume that the solar cells consist of a single segment. Outlier curves or segments, respectively, are rejected by only considering the largest clusters, with the additional constraint that the sizes of the used clusters are proportional to each other, and that not more than two different segments (as in Fig. 7c) can be expected in a cell. The number of rows and columns of a solar cell is determined by dividing the overall size of the curve grid by the estimated cell side lengths.

3.5.2 Curve grid outlier elimination

The estimated proportions are used to generate a synthetic planar grid that is registered against the curve grid intersections. Specifically, we use the rigid point set registration of Coherent Point Drift (CPD) [64] because it is deterministic and allows us to account for the proportion of outliers using a parameter $0 \leq w \leq 1$. We can immediately estimate w as

the fraction of points in the synthetic planar grid and the total number of intersections in the curve grid.

To ensure CPD convergence, initial positions of the synthetic planar grid should be sufficiently close to the curve grid intersections. We therefore estimate the translation and rotation of the planar grid to closely pre-align it with the grid we are registering against. The initial translation can be estimated as the curve grid intersection point closest to the image plane origin. The 2-D in-plane rotation is estimated from the average differences of two consecutive intersection points along each curve grid row and column. This results in two 2-D vectors which are approximately orthogonal to each other. The 2-D vector with the larger absolute angle is rotated by 90 % such that both vectors become roughly parallel. The estimated rotation is finally obtained as the average angle of both vectors.

3.5.3 Undistortion and rectification

The PV module configuration is used to undistort the whole image using Eq. (10). After eliminating the lens distortion, we use Direct Linear Transform (DLT) [38] to estimate the planar 2-D homography using the four corners of the curve grid with respect to the corners of the synthetic planar grid. The homography is used to remove perspective distortion from the undistorted curve grid.

The intersections of the perspective corrected curve grid may not align exactly with respect to the synthetic planar grid because individual solar cells are not always accurately placed in a perfect grid but rather with a margin of error. The remaining misalignment is therefore corrected via affine Moving Least Squares (MLS) [81], which warps the image using the planar grid intersections as control points distorted using the estimated lens parameters, and curve grid intersections are used as their target positions.

3.6 Estimation of the active solar cell area

We use solar cell images extracted from individual PV modules to generate a mask that represents the active solar cell area. Such masks allow to exclude the background and the busbars of a solar cell (see Fig. 8). In particular, active cell area masks are useful for detection of cell cracks since they allow to mask out the busbars, which can be incorrectly identified as cell cracks due to high similarity of their appearance [87,89].

Estimation of solar cell masks is related to the image labeling problem, where the goal is to classify every pixel into several predefined classes (in our case, the background and the active cell area). Existing approaches solve this problem using probabilistic graphical models, such as a Conditional Random Field (CRF) which learns the mapping in a supervised manner through contextual information [40]. However,

since the estimated curve grid already provides a global context, we tackle the pixelwise classification as a combination of adaptive thresholding and prior knowledge with regard to the straight shape of solar cells. Compared to CRFs, this approach does not require a training step and is easy to implement.

To this end, we use solar cells extracted from a PV module to compute a mean solar cell (see Figs. 8a, b). Since intensities within a mean solar cell image can exhibit a large range, we apply locally adaptive thresholding [68] on 25×25 pixels patches using their mean intensity, followed by a 15×15 morphological opening and flood filling to close any remaining holes. This leads to an initial binary mask.

Ragged edges at the contour are removed using vertical and horizontal cell profiles (Fig. 8b). The profiles are computed as pixelwise median of the initial mask along each image row or column, respectively. We combine the back-projection of these profiles with the convex hull of the binary mask determined with the method of Barber et al. [6] to account for cut-off corners using bitwise AND (cf., Fig. 8c). To further exclude repetitive patterns in the EL image of a solar cell, e.g., due to low passivation efficiency in the contact region (see Fig. 8d), we combine the initial binary mask and the augmented mask via bitwise XOR.

We note that solar cells are usually symmetric about both axes. Thus, the active solar cell area mask estimation can be restricted to only on quadrant of the average solar cell image to enforce mask symmetry. Additionally, the convex hull of the solar cell and its extra geometry can approximated by polygons [1] for a more compact representation.

3.7 Parameter tuning

The proposed solar cell segmentation pipeline relies on a set of hyperparameters that directly affect the segmentation robustness and accuracy. Table 1 provides an overview of all parameters with their values used in this work.

3.7.1 Manual search

Since the parameters of the proposed segmentation are intuitive and easily interpretable, it is straightforward to select them based on the setup used for EL image acquisition.

Main influence factors that must be considered when choosing the parameters are image resolution and physical properties of the camera lens.

Provided parameter values were found to work particularly well for high resolution EL images and standard camera lenses, as in our dataset (cf., Sect. 4.1). For low resolution EL images, however, the number of pyramid octaves and sub-levels will need to be increased to avoid missing important image details. Whereas, tensor voting proximity, on contrary, will need to be lowered, since the width of ridge edges in low

Table 1 Overview of segmentation pipeline parameters and their values used in this work

§	Symbol	Description	Used value
3.2.2	O	Number of octaves in Gaussian scale-space pyramid	5
	P	Number of sublevels in each octave	8
	σ	Gaussian scale-space standard deviation	1.6
	γ	Gaussian scale-space pyramid downsampling factor	2
3.2.3	ν	Tensor voting angular specificity	2
	$\varsigma_{1,2}$	Proximity of the 1 st and 2 nd tensor voting steps	15, 10
3.3.1		1-D sampling window for Gaussian-based subpixel position	21
3.3.2	ϑ	Maximum merge angle of two neighboring line segments	5°
3.3.3	ρ	Maximum error between fitted parabolic curve value at curve point	1.5
3.4.4	ϵ	Minimal change in error during refinement of lens distortion parameters	10 ⁻⁶
3.5		Solar cell aspect ratio	1 : 1
3.6		Locally adaptive thresholding window size	25 × 25

resolution images tends to be proportional to the image resolution. This immediately affects the size of the 1-D sampling window for determining the Gaussian-based subpixel position of curve points.

Curve extraction parameters correlate with the field-of-view of the EL camera lens. In particular for wide angle lenses, the merge angle ϑ must be increased.

Parabolic curve fit error ρ balances between robustness and accuracy of the segmentation result. The window size for locally adaptive thresholding used for estimation of solar cell masks correlates both with the resolution of EL images, but also with the amount of noise and texture variety in solar cells, e.g. due to cell cracks.

3.7.2 Automatic search

The parameters can also be automatically optimized in an efficient manner using random search [7,58,74,82,83,85] or Bayesian optimization [3,8,9,49,63,84] class of algorithms. Since this step involves supervision, pixelwise PV module annotations are needed. In certain cases, however, it may be not be possible to provide such annotations because individual defective PV cells can be hard to delineate, e.g., they appear completely dark. Also, the active solar cell area of defective cells is not always well-defined. Therefore, we refrained from automatically optimizing the hyperparameters in this work.

4 Evaluation

We evaluate the robustness and accuracy of our approach against manually annotated ground truth masks. Further, we compare the proposed approach against the method by

Sovetkin and Steland [86] on simplified masks, provide qualitative results and runtimes, and discuss limitations.

4.1 Dataset

We use a dataset consisting of 44 unique PV modules with various degrees of defects to manually select the parameters for the segmentation pipeline and validate the results. These images served as a reference during the development of the proposed method. The PV modules were captured in a testing laboratory setting at different orientations and using varying camera settings, such as exposure time. Some of EL images were post-processed by cropping, scaling, or rotation. This dataset consists of 26 monocrystalline and 18 polycrystalline solar cells. In total, these 44 solar modules consist of 2,624 solar cells out of which 715 are definitely defective with defects ranging from microcracks to completely disconnected cells and mechanically induced cracks (e.g., electrically insulated or conducting cracks, or cell cracks due to soldering [88]). 106 solar cells exhibit smaller defects that are not with certainty identifiable as completely defective, and 295 solar cells feature miscellaneous surface abnormalities that are no defects. The remaining 1,508 solar cells are categorized as functional without any perceivable surface abnormalities. The solar cells in imaged PV modules have a square aspect ratio (i.e., are quadratic).

The average resolution of the EL images is 2779.63×2087.35 pixels with a standard deviation of image width and height of 576.42 and 198.30 pixels, respectively. The median resolution is 3152×2046 pixels.

Additional eight test EL images (i.e., about 15% of the dataset) are used for the evaluation. Four modules are monocrystalline and the remaining four are polycrystalline. Their ground truth segmentation masks consist of hand-labeled solar cell segments. The ground truth additionally

specifies both the rows and columns of the solar cells, and their subdivisions. These images show various PV modules with a total of 408 solar cells. The resolution of the test EL images varies around $2649.50 \pm 643.20 \times 2074 \pm 339.12$ with a median image resolution of 2581.50×2046 .

Three out of four monocrystalline modules consist of 4×9 cells and the remaining monocrystalline module consists of 6×10 cells. All of their cells are subdivided by busbars into 3×1 segments.

The polycrystalline modules consist of 6×10 solar cells each. In two of the modules, every cell is subdivided into 3×1 segments. The cells of the other two modules are subdivided into 4×1 segments.

4.2 Evaluation metrics

We use two different metrics, pixelwise scores and the weighted Jaccard index to evaluate both the robustness and the accuracy of the proposed method and to compare our method against related work. In the latter case, we additionally use a third metric, the Root Mean Square Error (RMSE), to compute the segmentation error on simplified masks.

4.2.1 Root mean square error

The first performance metric is the RMSE given in pixels between the corners of the quadrilateral mask computed from the ground truth annotations and the corners estimated by the individual modalities. The metric provides a summary of the method's accuracy in absolute terms across all experiments.

4.2.2 Pixelwise classification

The second set of performance metrics are precision, recall, and the F_1 score [76]. These metrics are computed by considering cell segmentation as a multiclass pixelwise classification into background and active area of individual solar cells. A typical 60 cell PV module will therefore contain up to 61 class labels. A correctly segmented active area pixel is a true positive, the remaining quantities are defined accordingly. Pixelwise scores are computed globally with respect to all the pixels. Therefore, the differences between the individual results for these scores are naturally smaller than for metrics that are computed with respect to individual solar cells, such as the Jaccard index.

4.2.3 Weighted Jaccard Index

The third performance metric is the weighted Jaccard index [14,43], a variant of the metric widely known as Intersection-over-Union (IoU). This metric extends the common Jaccard index by an importance weighting of the input pixels. As the compared masks are not strictly binary either due to

antialiasing or interpolation during mask construction, we define importance of pixels by their intensity. Given two non-binary masks A and B , the weighted Jaccard similarity is

$$J_w = \frac{\sum_{\mathbf{u} \in \Omega} \min\{A(\mathbf{u}), B(\mathbf{u})\}}{\sum_{\mathbf{u} \in \Omega} \max\{A(\mathbf{u}), B(\mathbf{u})\}}. \quad (22)$$

The performance metric is computed on pairs of segmented cells and ground truth masks. A ground truth cell mask is matched to the segmented cell with the largest intersection area, thus taking structural coherence into account.

We additionally compute the Jaccard index of the background, which corresponds to the accuracy of the method to segment the whole solar module. Solar cell misalignment or missed cells will therefore penalize the segmentation accuracy to a high degree. Therefore, the solar module Jaccard index provides a summary of how well the segmentation performs per EL image.

4.3 Quantitative results

We evaluate the segmentation accuracy and the robustness of our approach using a fixed set of parameters as specified in Table 1 on EL images of PV modules acquired in a material testing laboratory.

4.3.1 Comparison to related work with simplified cell masks

The method by Sovetkin and Steland focuses on the estimation of the perspective transformation of the solar module and the extraction of solar cells. Radial distortion is corrected with a lens model of an external checkerboard calibration. The grid structure is fitted using a priori knowledge of the module topology. For this reason, we refer to the method as Perspective-corrected Grid Alignment (PGA). The method makes no specific proposal for mask generation and therefore yields rectangular solar cells.

In order to perform a comparison, the exact masks (cf., Fig. 9a) are restricted to quadrilateral shapes (cf., Fig. 9b). The quadrilateral mask is computed as the minimum circumscribing polygon with four sides, i.e., a quadrilateral, using the approach of Aggarwal et al. [1]. The quadrilateral exactly circumscribes the convex hull of the solar cell mask with all the quadrilateral sides flush to the convex hull.

PGA assumes that radial distortion is corrected by an external checkerboard calibration. This can be a limiting factor in practice. Hence, the comparison below considers both practical situations by running PGA on distorted images and on undistorted images using the distortion correction of this work.

Root Mean Square Error Table 2 provides the RMSE in pixels between the corners of the quadrilaterals computed by the respective modality and the quadrilateral mask estimated

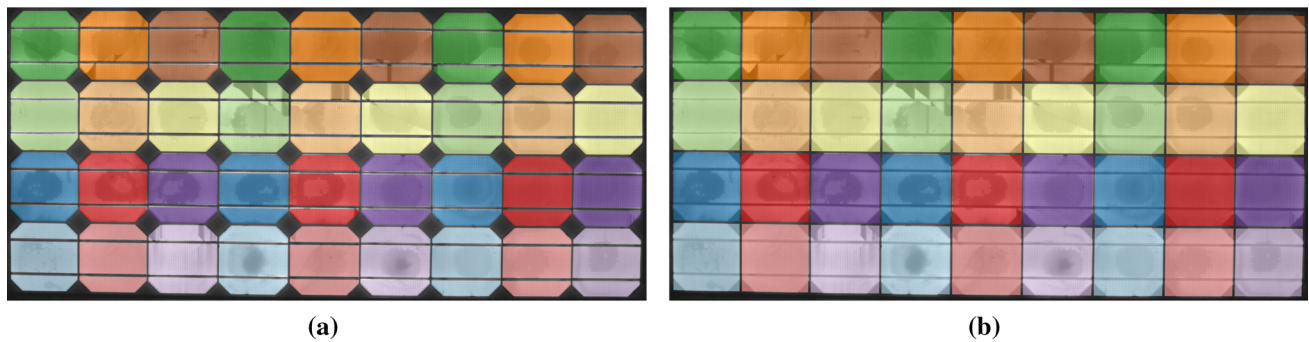


Fig. 9 Example of an exact mask (a) of solar cells estimated using the proposed approach and a quadrilateral mask (b) determined from the exact mask. The latter is used for comparison against the method of Sovetkin and Steland [86]. Both masks are shown as color overlays. Different colors denote different instances of solar cells

Table 2 Root Mean Square Error (RMSE), in pixels, of the distance between the corners of the quadrilateral mask determined from the ground truth annotations and the corners determined by the respective method in all eight test images

	Distorted				Undistorted			
	PGA	Proposed			PGA	Proposed		
		ω	(ω, s_x, c)	w/ MLS		ω	(ω, s_x, c)	w/ MLS
Monocrystalline	6.09	2.71	2.61	2.64	4.00	2.88	2.68	2.61
Polycrystalline	5.32	2.56	2.52	2.45	2.76	2.33	1.91	1.77
Overall	5.65	2.62	2.56	2.53	3.32	2.55	2.26	2.15

Bold face denotes smallest error

from the ground truth. The metric is provided for monocrystalline and polycrystalline solar wafers separately, and for both types combined. In all cases, the proposed approach outperforms both PGA variants. We particularly notice that PGA greatly benefits from lens distortion estimation. This underlines our observation that the latter is essential for highly accurate segmentation.

Pixelwise Classification Pixelwise scores for the simplified masks of both methods are given in Table 3. For monocrystalline PV modules, PGA generally achieves higher scores. However, highest scores are achieved only for images for which the lens distortion has been removed. The proposed method fails to segment a row of cells in a solar module resulting in a lower recall. However, for polycrystalline PV modules, the proposed method consistently outperforms PGA. In the overall score, the proposed method also outperforms the best-case evaluation for PGA on undistorted images. However, PGA has highest recall, which is due to the lower number of parameters of PGA.

Weighted Jaccard Index The Jaccard scores summarized as boxplots in Fig. 10 support the pixelwise classification scores, showing that the proposed method is more accurate than PGA. The latter, however, is slightly more robust. For complete modules, the considerable spread of PGA is partially attributed to one major outlier. Overall, the proposed segmentation pipeline is highly accurate. Particularly once a cell is detected, the cell outline is accurately and robustly segmented.

4.3.2 Ablation study

We ablate the lens distortion parameters and the post hoc application of affine MLS to investigate their effect on the accuracy and the success rate of the segmentation process. The ablation is performed both on original (i.e., distorted) EL images and undistorted ones.

Distorted vs. Undistorted EL Images For the ablation study, we consider two main cases. In the undistorted case, both reference and predicted masks are unwarped using estimated lens distortion parameters. Then, quadrilaterals are fitted to individual cell masks to allow a comparison against PGA which always yields such quadrilateral cell masks. For a fair comparison, PGA is also applied to undistorted images.

In the distorted case, however, the comparison is performed in the original image space. Since the proposed method yields a curved grid after applying the inverse of lens distortion, we synthesize a regular grid from backwarped cell masks. Specifically, we extract the contours of estimated solar cell masks to obtain the coordinates of the quadrilateral in the unwarped image, and then apply the inverse of estimated geometric transformations to rectangle coordinates. Afterwards, we fit lines to each side of the backwarped quadrilaterals along grid rows and columns. From their intersections we finally obtain the corner coordinates of each solar cells in the distorted image which we can use for comparison against distorted PGA results.

Parameterization First, we reduce the lens distortion model to a single radial distortion parameter ω and assume both

Table 3 Pixelwise classification scores for quadrilateral masks estimated using PGA and the proposed approach

	Distorted				Undistorted			
	PGA	Proposed			PGA	Proposed		
		ω	$(\omega, s_x, \mathbf{c})$	w/ MLS		ω	$(\omega, s_x, \mathbf{c})$	w/ MLS
(a) Monocrystalline								
Precision	97.55%	97.77%	99.18%	98.98%	98.43%	98.34%	99.67%	99.53%
Recall	98.37%	82.08%	97.53%	97.71%	98.87%	81.56%	96.94%	97.14%
F_1 score	97.95%	86.57%	98.24%	98.24%	98.65%	86.46%	98.18%	98.21%
Accuracy	98.19%	88.43%	98.34%	98.32%	98.82%	88.55%	98.33%	98.38%
(b) Polycrystalline								
Precision	97.22%	97.70%	98.82%	98.77%	98.36%	97.40%	99.52%	99.59%
Recall	97.70%	88.35%	99.29%	99.36%	99.29%	87.08%	99.17%	99.35%
F_1 score	97.45%	91.32%	99.05%	99.06%	98.82%	90.42%	99.35%	99.47%
Accuracy	97.13%	92.44%	98.89%	98.90%	98.66%	92.06%	99.25%	99.39%
(c) Overall								
Precision	97.37%	97.30%	99.00%	98.88%	98.38%	97.09%	99.58%	99.56%
Recall	97.78%	82.36%	98.27%	98.39%	99.01%	81.15%	97.97%	98.18%
F_1 score	97.57%	87.45%	98.60%	98.60%	98.69%	86.60%	98.73%	98.83%
Accuracy	97.74%	90.15%	98.58%	98.57%	98.75%	90.06%	98.72%	98.81%

Bold face denotes the best performing method

PGA (distorted) PGA (undistorted) Proposed

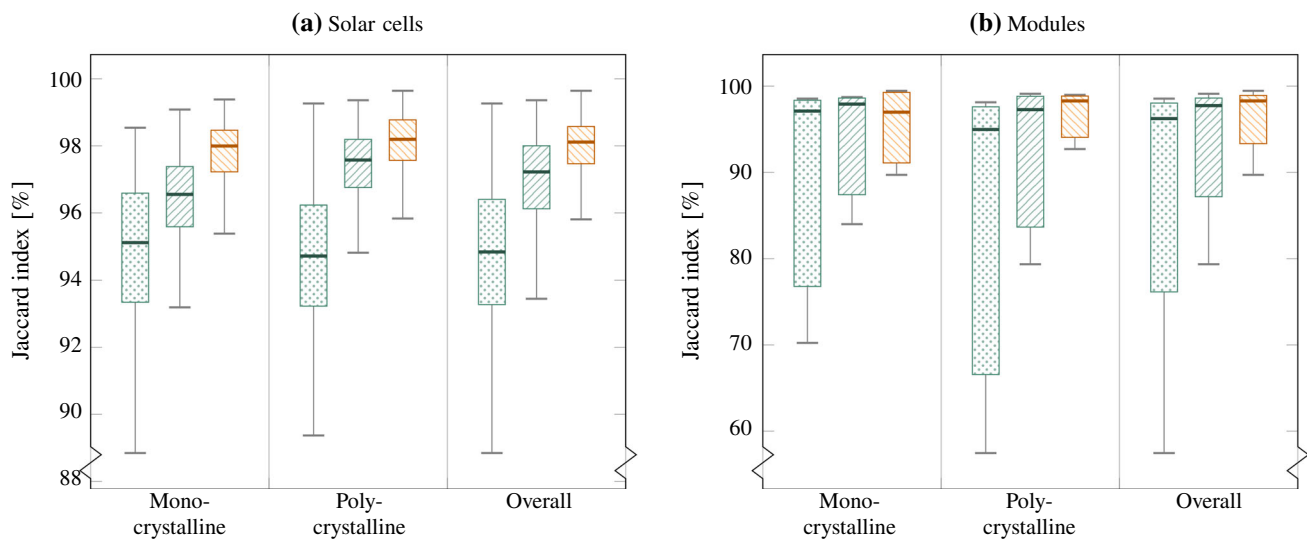


Fig. 10 Boxplots of Jaccard scores for the three evaluated modalities. The Jaccard scores are computed against hand-labeled ground truth masks. In (a), the scores are computed for the individual solar cells. In (b), the scores are evaluated against the whole solar modules. The two left-most groups in each figure correspond to boxplots with respect to different solar wafers. Whereas the right-most group summarizes the performance of both solar wafer types combined

square aspect ratio (i.e., $s_x = 1$) and the center of distortion to be located in the image center. During optimization, these two parameters are kept constant. In this experiment, we also do not correct the curve grid using affine MLS. The comparison against PGA shows that such a simplistic lens model is still more accurate than PGA both in the distorted and undistorted cases (cf., Table 2). However, while the precision is

high, the recall and therefore the F_1 score drops considerably (cf., Table 3). The reason for this is that such a lens parametrization is too rigid. As a consequence, this weakens the grid detection: correctly detected curves are erroneously discarded because of inaccuracies of the lens model with only a single parameter ω instead of four parameters $(\omega, s_x, \mathbf{c})$.

Table 4 Pixelwise classification scores for exact masks estimated using the proposed approach

Metric	Monocrystalline	Polycrystalline	Overall
Precision	97.47%	97.49%	97.53%
Recall	96.93%	98.90%	97.77%
F_1 score	97.09%	98.19%	97.62%
Accuracy	97.67%	97.97%	97.80%

For the next comparison, we increase the number of degrees of freedom. Both the distortion aspect ratio s_x and the center of distortion \mathbf{c} are refined in addition to the radial distortion parameter ω . Curve grid correction via affine MLS is again omitted. This parametrization achieves much improved RMSE and segmentation success rates.

Finally, we use the full parametrization, i.e., we refine all lens distortion parameters (ω , s_x , \mathbf{c}) and apply post hoc correction via affine MLS. This model is denoted as *w/MLS*. *Discussion* We summarize the results of the ablation study in Tables 2 and 3. Here, *w/MLS* denotes the full model that includes the correction step via affine MLS. The full model with post hoc affine MLS grid correction performs in many instances best. However, applying MLS is not always beneficial. Particularly, for monocrystalline PV modules, grid correction does not always improve the results.

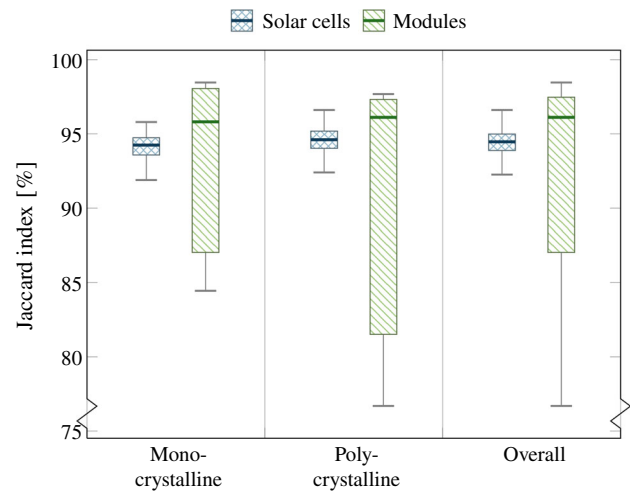
We conclude that the proposed joint lens model estimation with full parametrization and grid detection is essential for robustness and accuracy of the segmentation. Since the subsequent grid correction using affine MLS only marginally improves the results, its application can be seen as optional.

4.3.3 Segmentation performance with exact cell masks

To allow an exact comparison of the segmentation results to the ground truth, we inverse-warp the estimated solar cell masks back to the original image space by using the determined perspective projection and lens distortion parameters. This way, the estimated solar module masks will as exactly as possible overlay the hand-labeled ground truth masks.

Pixelwise Classification Table 4 summarizes the pixelwise classification scores for the exact masks estimated using the proposed method. The method is more robust on polycrystalline PV modules than on monocrystalline modules. However, for both module types, the method achieves a very high overall accuracy beyond 97 % for all metrics. Investigation of failure cases for monocrystalline modules reveals difficulties on cells where large gaps coincide with cell cracks and ragged edges.

Weighted Jaccard Index Jaccard scores for exact masks are given in Fig. 11. The scores confirm the results of the pixelwise metrics. Notably, the interquartile range (IQR) of individual cells has a very small spread, which indicates

**Fig. 11** Boxplots of Jaccard scores for the proposed approach

a highly consistent segmentation. The IQR of whole modules is slightly larger. This is, however, not surprising since the boxplots summarize the joint segmentation scores across multiple modules.

4.4 Qualitative results

Figure 12 shows the qualitative results of the segmentation pipeline on four test images. The two results in the left column are computed on monocrystalline modules, the two results in the right column on polycrystalline modules. The estimated solar module curve grids are highly accurate. Even in presence of complex texture intrinsic to the material, the accuracy of the predicted solar module curve grid is not affected.

4.5 Runtime evaluation

Figure 13 breaks down the average time taken by the individual steps of the segmentation pipeline. Figure 14 summarizes the contribution of individual pipeline steps to the overall processing time for all 44 images. The timings were obtained on a consumer system with an Intel i7-3770K CPU clocked at 3.50 GHz and 32 GB of RAM. The first three stages of the segmentation pipeline are implemented in C++ whereas the last stage (except for MLS image deformation) is implemented in Python.

For this benchmark, EL images were processed sequentially running only on the CPU. Note, however, that the implementation was not optimized in terms of the runtime and only parts of the pipeline utilize all available CPU cores. To this end, additional speedup can be achieved by running parts of the pipeline in parallel or even on a GPU.

On average, it takes 1 min and 6 s to segment all solar cells in a high resolution EL image (cf., Fig. 14). Preprocessing is computationally most expensive, curve and cell extraction

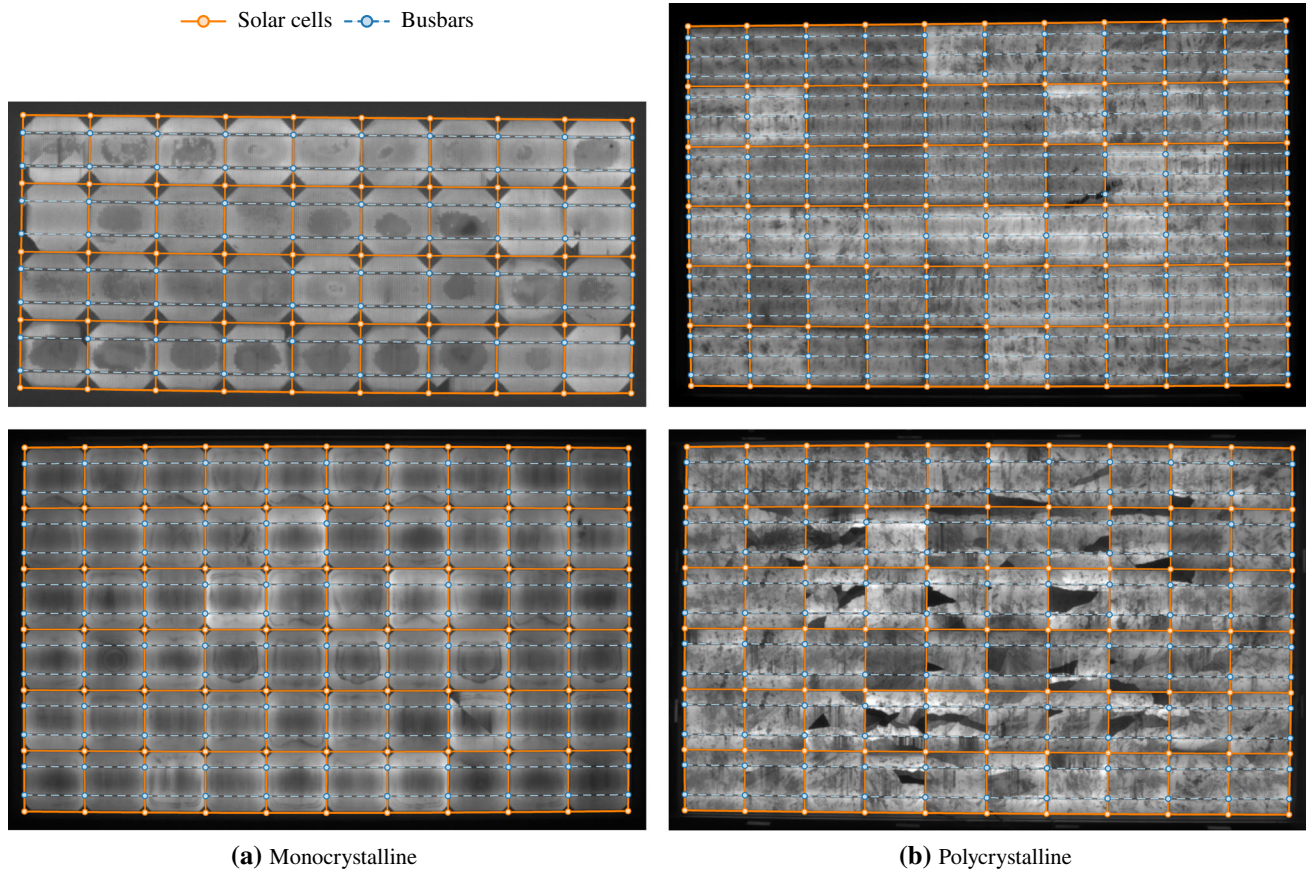


Fig. 12 Qualitative segmentation results of four test images depicting the estimated curve grid superimposed over the contrast-normalized input EL image. For visualization purposes, the original EL images were cropped

are on average cheapest. The standard deviation of the model estimation step is highest (see Fig. 13), which is mostly due to dependency upon the total number of ridge edges and the number of resulting curves combined with the probabilistic nature of LO-RANSAC.

Interestingly, processing EL images of monocrystalline solar modules takes slightly longer on average than processing polycrystalline solar modules. This is due to large gaps between ridges caused by cut-off corners that produce many disconnected curve segments which must be merged first. Conversely, curve segments in polycrystalline solar modules are closer, which makes it more likely that several curve segments are combined early on.

An average processing time of 1 min and 6 s is substantially faster than manual processing, which takes at least several minutes. For on-site EL measurements with in-situ imaging of PV modules, the processing times must be further optimized, likely by at least a factor of ten. However, in other imaging environments, for example material testing laboratories, the runtime is fully sufficient, given that the handling of each module for EL measurements and the per-

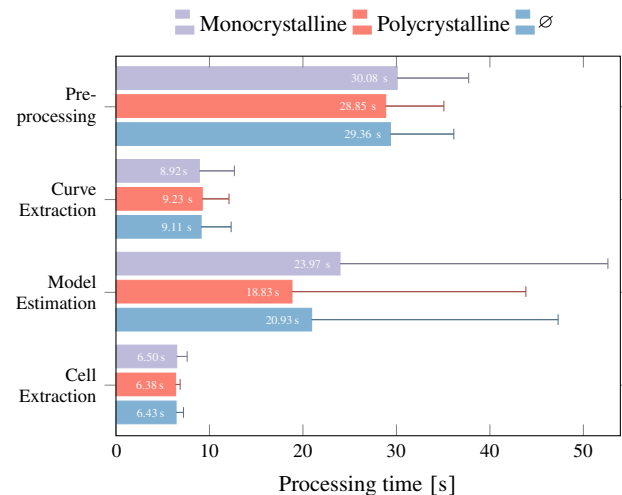


Fig. 13 Average time taken by individual steps of the segmentation pipeline, in seconds. The error bars denote the upper range of the standard deviation

formance evaluation impose much more severe scheduling bottlenecks.

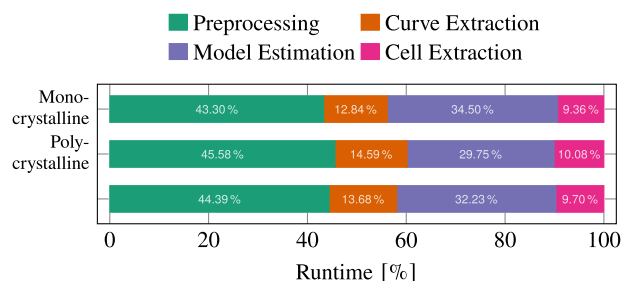


Fig. 14 Relative contribution of the average processing time for individual pipeline steps to the overall runtime with respect to different solar module types and both types combined

4.6 Limitations

Mounts that hold PV modules may cause spurious ridge edges. Early stages of the segmentation focus on ridges without analyzing the whole image content, which may occasionally lead to spurious edges and eventually to an incorrect segmentation. Therefore, automatic image cropping prior to PV module segmentation could help reduce segmentation failures due to visible mounts.

While the algorithm is able to process disconnected (dark) cells, rows or columns with more than 50 % of disconnected cells pose a difficulty in correctly detecting the grid due to insufficient edge information. However, we observed that also human experts have problems to determine the contours under such circumstances.

We also observed that smooth edges can result in segmentation failures. This is because the stickiness of smooth edges is weak and may completely fade away after non-maximum suppression. This problem is also related to situations where the inter-cell borders are exceptionally wide. In such cases, it is necessary to adjust the parameters of the ridgeness filter and the proximity of the tensor voting.

5 Conclusions

In this work, we presented a fully automatic segmentation method for precise extraction of solar cells from high resolution EL images. The proposed segmentation is robust to underexposure, and works robustly in presence of severe defects on solar cells. This can be attributed to the proposed preprocessing and the ridgeness filtering, coupled with tensor voting to robustly determine the inter-cell borders and bus-bars. The segmentation is highly accurate, which allows to use its output for further inspection tasks, such as automatic classification of defective solar cells and the prediction of power loss.

We evaluated the segmentation with the Jaccard index on eight different PV modules consisting of 408 hand-labeled solar cells. The proposed approach is able to segment solar

cells with an accuracy of 97.80 %. With respect to classification performance, the segmentation pipeline reaches an F_1 score of 97.61 %.

Additionally, we compared the proposed method against the PV module detection approach by Sovetkin and Steland [86], which is slightly more robust but less accurate than our method. The comparison also shows that our joint lens distortion estimation and grid detection approach achieves a higher accuracy than a method that decouples both steps.

Beyond the proposed applications, the method can serve as a starting point for bootstrapping deep learning architectures that could be trained end-to-end to directly segment the solar cells. Future work may include to investigate the required adaptations and geometric relaxations for using the method not only in manufacturing setting but also in the field. Such relaxations could be achieved, for instance, by performing the grid detection end-to-end using a CNN.

Given that grid structure is pervasive in many different problem domains, the proposed joint lens estimation and grid identification may also find other application fields, for example the detection of PV modules in aerial imagery of solar power plants, building facade segmentation, and checker-board pattern detection for camera calibration.

Acknowledgements This work was funded by Energy Campus Nuremberg (EnCN) and partially supported by the Research Training Group 1773 “Heterogeneous Image Systems” funded by the German Research Foundation (DFG).

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Aggarwal, A., Chang, J.S., Yap, C.K.: Minimum area circumscribing polygons. *Vis. Comput.* **1**(2), 112–117 (1985). <https://doi.org/10.1007/BF01898354>
2. Ahmed, M., Farag, A.: Nonmetric calibration of camera lens distortion: differential methods and robust estimation. *IEEE Trans. Image Process.* **14**(8), 1215–1230 (2005). <https://doi.org/10.1109/TIP.2005.846025>
3. Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M.: Optuna: a next-generation hyperparameter optimization framework. In: *Proceedings of the 25th ACM SIGKDD International Conference on*

- Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA, KDD '19, pp. 2623–2631, (2019) <https://doi.org/10.1145/3292500.3330701>
4. Aloimonos, J.: Shape from texture. *Biol. Cybern.* **58**(5), 345–360 (1988). <https://doi.org/10.1007/BF00363944>
 5. Anwar, S.A., Abdullah, M.Z.: Micro-crack detection of multicrystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique. *EURASIP J. Image Video Process.* **2014**(1), 15 (2014). <https://doi.org/10.1186/1687-5281-2014-15>
 6. Barber, C.B., Dobkin, D.P., Dobkin, D.P., Huhdanpaa, H.: The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.* **22**(4), 469–483 (1996). <https://doi.org/10.1145/235815.235821>
 7. Bergstra, J., Bengio, Y.: Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13**, 281–305 (2012)
 8. Bergstra, J., Bardenet, R., Bengio, Y., Kégl, B.: Algorithms for hyper-parameter optimization. In: Shawe-Taylor, J., Zemel, R., Bartlett, P., Pereira, F., Weinberger, K.Q. (eds.) *Advances in Neural Information Processing Systems*, vol. 24, pp. 2546–2554. Curran Associates, New York (2011)
 9. Bergstra, J., Yamins, D., Cox, D.D.: Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. In: *Proceedings of the 30th International Conference on International Conference on Machine Learning, ICML'13*, pp. 115–123 (2013)
 10. Breitenstein, O., Bauer, J., Bothe, K., Hinken, D., Müller, J., Kwapił, W., Schubert, M.C., Warta, W.: Can luminescence imaging replace lock-in thermography on solar cells? *IEEE J. Photovolt.* **1**(2), 159–167 (2011). <https://doi.org/10.1109/JPHOTOV.2011.2169394>
 11. Brown, D.C.: Close-range camera calibration. *Photogramm. Eng. Remote Sens.* **37**, 855–866 (1971). <https://doi.org/10.1.1.14.6358>
 12. Buerhop-Lutz, C., Deutsch, S., Maier, A., Gallwitz, F., Berger, S., Doll, B., Hauch, J., Camus, C., Brabec, C.J.: A benchmark for visual identification of defective solar cells in electroluminescence imagery. In: *35th European PV Solar Energy Conference and Exhibition*, pp. 1287–1289 (2018) <https://doi.org/10.4229/35thEUPVSEC20182018-5CV.3.15>
 13. Chenni, R., Makhlof, M., Kerbache, T., Bouzid, A.: A detailed modeling method for photovoltaic cells. *Energy* **32**(9), 1724–1730 (2007). <https://doi.org/10.1016/j.energy.2006.12.006>
 14. Chierichetti, F., Kumar, R., Pandey, S., Vassilvitskii, S.: Finding the Jaccard median. In: *Symposium on Discrete Algorithms*, pp. 293–311, Austin, Texas (2010)
 15. Chum, O., Matas, J., Kittler, J.: Locally optimized RANSAC. In: Michaelis, B., Krell, G. (eds.) *Pattern Recognition*, vol. 2781, pp. 236–243. Springer, Berlin (2003). https://doi.org/10.1007/978-3-540-45243-0_31
 16. Claus, D., Fitzgibbon, A.W.: A plumbline constraint for the rational function lens distortion model. In: *British Machine Vision Conference (BMVC)*, pp. 99–108 (2005) <https://doi.org/10.5244/C.19.10>
 17. Claus, D., Fitzgibbon, A.W.: A rational function lens distortion model for general cameras. *Conf. Comput. Vis. Pattern Recognit. (CVPR)* **1**, 213–219 (2005b). <https://doi.org/10.1109/CVPR.2005.43>
 18. Deutsch, S., Christlein, V., Berger, S., Buerhop-Lutz, C., Maier, A., Gallwitz, F., Riess, C.: Automatic classification of defective photovoltaic module cells in electroluminescence images. *Sol. Energy* **185**, 455–468 (2019). <https://doi.org/10.1016/j.solener.2019.02.067>. (arXiv:1807.02894)
 19. Devernay, F.: A non-maxima suppression method for edge detection with sub-pixel accuracy. Technical Report, RR-2724, INRIA (1995) <https://hal.inria.fr/inria-00073970>
 20. Devernay, F., Faugeras, O.: Straight lines have to be straight. *Mach. Vis. Appl.* **13**(1), 14–24 (2001). <https://doi.org/10.1007/PL00013269>
 21. Douglas, D.H., Peucker, T.K.: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartogr. Int. J. Geogr. Inf. Geovis.* **10**(2), 112–122 (1973). <https://doi.org/10.3138/fm57-6770-u75u-7727>
 22. Ester, M., Kriegel, H.P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, AAAI Press, Portland, OR, USA, KDD'96*, pp. 226–231 (1996)
 23. Fabijańska, A.: A survey of subpixel edge detection methods for images of heat-emitting metal specimens. *Int. J. Appl. Math. Comput. Sci.* **22**(3), 695–710 (2012). <https://doi.org/10.2478/v10006-012-0052-3>
 24. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* **24**(6), 381–395 (1981). <https://doi.org/10.1145/358669.358692>
 25. Fitzgibbon, A.: Simultaneous linear estimation of multiple view geometry and lens distortion. *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **1**, 125–132 (2001). <https://doi.org/10.1109/CVPR.2001.990465>
 26. Flocke, N.: Algorithm 954: an accurate and efficient cubic and quartic equation solver for physical applications. *ACM Trans. Math. Softw.* **41**(4), 30:1–30:24 (2015). <https://doi.org/10.1145/2699468>
 27. Franken, E., van Almsick, M., Rongen, P., Florack, L., ter Haar Romeny, B.: An efficient method for tensor voting using steerable filters. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *European Conference on Computer Vision (ECCV)*, pp. 228–240 (2006) https://doi.org/10.1007/11744085_18
 28. Franken, E., Rongen, P., van Almsick, M., ter Haar Romeny, B.: Detection of electrophysiology catheters in noisy fluoroscopy images. In: Larsen, R., Nielsen, M., Sparring, J. (eds.) *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2006*, Springer, Berlin, Heidelberg, pp. 25–32 (2006) https://doi.org/10.1007/11866763_4
 29. Fürsattel, P., Dotenco, S., Placht, S., Balda, M., Maier, A., Riess, C.: OCPAD—occluded checkerboard pattern detector. In: *Winter Conference on Applications of Computer Vision (WACV)*, IEEE, (2016) <https://doi.org/10.1109/WACV.2016.7477565>
 30. Grompone von Gioi, R., Randall, G.: A sub-pixel edge detector: an implementation of the Canny/Devernay algorithm. *Image Process. On Line* **7**, 347–372 (2017). <https://doi.org/10.5201/ipol.2017.216>
 31. Girshick, R.B.: Fast R-CNN. In: *IEEE International Conference on Computer Vision*, pp. 1440–1448 (2015). <https://doi.org/10.1109/ICCV.2015.169>
 32. Girshick, R.B., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587 (2014). <https://doi.org/10.1109/CVPR.2014.81>
 33. Golub, G.H., Van Loan, C.F.: *Matrix Computations*, 4th edn. Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press (2013)
 34. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 4th edn. Pearson, New York (2018)
 35. Goodfellow, I., Bengio, Y., Courville, A.: *Deep Learning*. MIT Press, Cambridge (2016)
 36. Ha, H., Perdoch, M., Alismail, H., Kweon, I.S., Sheikh, Y.: Deltille grids for geometric camera calibration. In: *International Conference on Computer Vision (ICCV)*, pp. 5354–5362 (2017). <https://doi.org/10.1109/ICCV.2017.571>
 37. Harker, M., O'Leary, P., Zsombor-Murray, P.: Direct type-specific conic fitting and eigenvalue bias correction. *Image Vis. Comput.* **26**(3), 372–381 (2008). <https://doi.org/10.1016/j.imavis.2006.12.006>
 38. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd edn. Cambridge University Press, New York (2004)

39. He, K., Gkioxari, G., Dollár, P., Girshick, R.B.: Mask R-CNN. In: IEEE International Conference on Computer Vision, pp. 2980–2988 (2017). <https://doi.org/10.1109/ICCV.2017.322>
40. He, X., Zemel, R.S., Carreira-Perpinan, M.A.: Multiscale conditional random fields for image labeling. IEEE Conf. Comput. Vis. Pattern Recognit. **2**, 695–702 (2004). <https://doi.org/10.1109/CVPR.2004.1315232>
41. Hoffmann, M., Ernst, A., Bergen, T., Hettenger, S., Garbas, J.U.: A robust chessboard detector for geometric camera calibration. In: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP), pp. 34–43 (2017). <https://doi.org/10.5220/0006104300340043>
42. IEC TS 60904-13:2018 (2018) Photovoltaic devices—part 13: Electroluminescence of photovoltaic modules. Technical specification, International Electrotechnical Commission
43. Ioffe, S.: Improved consistent sampling, weighted minhash and L^1 sketching. In: International Conference on Data Mining, pp. 246–255 (2010). <https://doi.org/10.1109/ICDM.2010.80>
44. Jaderberg, M., Simonyan, K., Zisserman, A., Kavukcuoglu, K.: Spatial transformer networks. In: Cortes, C., Lawrence, N.D., Lee, D.D., Sugiyama, M., Garnett, R. (eds.) Advances in Neural Information Processing Systems, 28, pp. 2017–2025. Curran Associates Inc, New York (2015)
45. Jenkins, M.A., Traub, J.F.: A three-stage algorithm for real polynomials using quadratic iteration. J. Numer. Anal. **7**(4), 545–566 (1970). <https://doi.org/10.1137/0707045>
46. Karatepe, E., Boztepe, M., Çolak, M.: Development of a suitable model for characterizing photovoltaic arrays with shaded solar cells. Sol. Energy **81**(8), 977–992 (2007). <https://doi.org/10.1016/j.solener.2006.12.001>
47. Kaushika, N.D., Gautam, N.K.: Energy yield simulations of inter-connected solar PV arrays. IEEE Trans. Energy Convers. **18**(1), 127–134 (2003). <https://doi.org/10.1109/TEC.2002.805204>
48. Kovese, P.: MATLAB and Octave functions for computer vision and image processing (2017). <http://www.peterkovese.com/matlabfns>
49. Kushner, H.J.: A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. J. Basic Eng. **86**(1), 97–106 (1964). <https://doi.org/10.1115/1.3653121>
50. Levenberg, K.: A method for the solution of certain non-linear problems in least squares. Q. Appl. Math. **2**(2), 164–168 (1944)
51. Li, Y., Qi, H., Dai, J., Ji, X., Wei, Y.: Fully convolutional instance-aware semantic segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 4438–4446 (2017). <https://doi.org/10.1109/CVPR.2017.472>
52. Lin, C.H., Lucey, S.: Inverse compositional spatial transformer networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, pp. 2252–2260 (2017). <https://doi.org/10.1109/CVPR.2017.242>
53. Lindeberg, T.: Scale-space theory: a basic tool for analyzing structures at different scales. J. Appl. Stat. **21**(1–2), 225–270 (1994). <https://doi.org/10.1080/757582976>
54. Lindeberg, T.: Edge detection and ridge detection with automatic scale selection. In: Proceedings of the Conference on Computer Vision and Pattern Recognition, pp. 465–470 (1996). <https://doi.org/10.1109/CVPR.1996.517113>
55. Lindeberg, T.: Edge detection and ridge detection with automatic scale selection. Int. J. Comput. Vis. **30**(2), 117–156 (1998). <https://doi.org/10.1023/A:1008097225773>
56. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3431–3440 (2015). <https://doi.org/10.1109/CVPR.2015.7298965>
57. Marquardt, D.W.: An algorithm for least-squares estimation of non-linear parameters. J. Soc. Ind. Appl. Math. **11**(2), 431–441 (1963). <https://doi.org/10.1137/0111030>
58. Masri, S.F., Bekey, G.A., Safford, F.B.: A global optimization algorithm using adaptive random search. Appl. Math. Comput. **7**(4), 353–375 (1980). [https://doi.org/10.1016/0096-3003\(80\)90027-2](https://doi.org/10.1016/0096-3003(80)90027-2)
59. Mauk, M.G.: Image processing for solar cell analysis, diagnostics and quality assurance inspection. In: Image Processing: Concepts, Methodologies, Tools, and Applications, IGI Global, chap 71, pp. 1426–1462 (2013)
60. Mayr, M., Hoffmann, M., Maier, A., Christlein, V.: Weakly supervised segmentation of cracks on solar cells using normalized L_p norm. In: 2019 IEEE International Conference on Image Processing (ICIP), pp. 1885–1889 (2019). <https://doi.org/10.1109/ICIP.2019.8803116>
61. Medioni, G., Tang, C.K., Lee, M.S.: Tensor voting: theory and applications. In: Proceedings of RFIA (2000)
62. Mehta, S., Azad, A.P., Chennamgath, S.A., Raykar, V., Kalyanaraman, S.: DeepSolarEye: power loss prediction and weakly supervised soiling localization via fully convolutional networks for solar panels. In: Winter Conference on Applications of Computer Vision (WACV), pp. 333–342 (2018). <https://doi.org/10.1109/WACV.2018.00043>
63. Močkus, J.: On Bayesian methods for seeking the extremum. In: Marchuk, G.I. (ed.) IFIP Technical Conference on Optimization Techniques, pp. 400–404. Springer, Berlin (1975)
64. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. **32**(12), 2262–2275 (2010). <https://doi.org/10.1109/TPAMI.2010.46>
65. Nian, B., Fu, Z., Wang, L., Cao, X.: Automatic detection of defects in solar modules: image processing in detecting. In: 2010 International Conference on Computational Intelligence and Software Engineering, IEEE, pp. 1–4 (2010). <https://doi.org/10.1109/WICOM.2010.5600703>
66. Nocedal, J., Wright, S.J.: Numerical Optimization, Operations Research and Financial Engineering, vol. 43, 2nd edn. Springer, New York (2006). <https://doi.org/10.1007/978-0-387-40065-5>
67. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. **9**(1), 62–66 (1979). <https://doi.org/10.1109/TSMC.1979.4310076>
68. Pitas, I.: Digital Image Processing Algorithms. Prentice-Hall Inc, Upper Saddle River (1993)
69. Placht, S., Fürsattel, P., Mengue, E.A., Hofmann, H., Schaller, C., Balda, M., Angelopoulou, E.: ROCHADE: robust checkerboard advanced detection for camera calibration. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) European Conference on Computer Vision (ECCV), Lecture Notes in Computer Science, vol. 8692, pp. 766–779 (2014). https://doi.org/10.1007/978-3-319-10593-2_50
70. Potthoff, T., Bothe, K., Eitner, U., Hinken, D., Köntges, M.: Detection of the voltage distribution in photovoltaic modules by electroluminescence imaging. Prog. Photovolt. Res. Appl. **18**(2), 100–106 (2010). <https://doi.org/10.1002/ppp.941>
71. Prasad, D.K., Leung, M.K.H., Quek, C., Cho, S.Y.: A novel framework for making dominant point detection methods non-parametric. Image Vis. Comput. **30**(11), 843–859 (2012). <https://doi.org/10.1016/j.imavis.2012.06.010>
72. Quaschnig, V., Hanitsch, R.: Numerical simulation of current-voltage characteristics of photovoltaic systems with shaded solar cells. Sol. Energy **56**(6), 513–520 (1996). [https://doi.org/10.1016/0038-092X\(96\)00006-0](https://doi.org/10.1016/0038-092X(96)00006-0)
73. Ramer, U.: An iterative procedure for the polygonal approximation of plane curves. Comput. Graph. Image Process. **1**(3), 244–256 (1972). [https://doi.org/10.1016/S0146-664X\(72\)80017-0](https://doi.org/10.1016/S0146-664X(72)80017-0)
74. Rastrigin, L.A.: The convergence of the random search method in the extremal control of a many parameter system. Autom. Remote Control **24**, 1337–1342 (1963)
75. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans.

- Pattern Anal. Mach. Intell. **39**(6), 1137–1149 (2017). <https://doi.org/10.1109/TPAMI.2016.2577031>
76. van Rijsbergen, C.J.: Information Retrieval, 2nd edn. Butterworth-Heinemann, Oxford (1979)
 77. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
 78. Rosten, E., Loveland, R.: Camera distortion self-calibration using the plumb-line constraint and minimal Hough entropy. Mach. Vis. Appl. **22**(1), 77–85 (2011). <https://doi.org/10.1007/s00138-009-0196-9>. (arXiv:0810.4426)
 79. Rufli, M., Scaramuzza, D., Siegwart, R.: Automatic detection of checkerboards on blurred and distorted images. In: International Conference on Intelligent Robots and Systems, pp. 3121–3126 (2008). <https://doi.org/10.1109/IROS.2008.4650703>
 80. Saeed, K., Tabędzki, M., Rybnik, M., Adamski, M.: K3M: a universal algorithm for image skeletonization and a review of thinning techniques. Int. J. Appl. Math. Comput. Sci. **20**(2), 317–335 (2010). <https://doi.org/10.2478/v10006-010-0024-4>
 81. Schaefer, S., McPhail, T., Warren, J.: Image deformation using moving least squares. ACM Trans. Graph. **25**(3), 533 (2006). <https://doi.org/10.1145/1141911.1141920>
 82. Schrack, G., Choit, M.: Optimized relative step size random searches. Math. Program. **10**(1), 230–244 (1976). <https://doi.org/10.1007/BF01580669>
 83. Schumer, M., Steiglitz, K.: Adaptive step size random search. IEEE Trans. Autom. Control **13**(3), 270–276 (1968). <https://doi.org/10.1109/TAC.1968.1098903>
 84. Snoek, J., Larochelle, H., Adams, R.P.: Practical Bayesian optimization of machine learning algorithms. In: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems, vol. 25, pp. 2951–2959. Curran Associates Inc, New York (2012)
 85. Solis, F.J., Wets, R.J.B.: Minimization by random search techniques. Math. Oper. Res. **6**(1), 19–30 (1981). <https://doi.org/10.1287/moor.6.1.19>
 86. Sovetkin, E., Steland, A.: Automatic processing and solar cell detection in photovoltaic electroluminescence images. Integr. Comput. Aided Eng. **26**(2), 123–137 (2019). <https://doi.org/10.3233/ICA-180588>
 87. Spataru, S., Hacke, P., Sera, D.: Automatic detection and evaluation of solar cell micro-cracks in electroluminescence images using matched filters. In: 2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC), pp. 1602–1607 (2016). <https://doi.org/10.1109/PVSC.2016.7749891>
 88. Spertino, F., Ciocia, A., Leo, P.D., Tommasini, R., Berardone, I., Corrado, M., Infuso, A., Paggi, M.: A power and energy procedure in operating photovoltaic systems to quantify the losses according to the causes. Sol. Energy **118**, 313–326 (2015). <https://doi.org/10.1016/j.solener.2015.05.033>
 89. Stromer, D., Vetter, A., Oezkan, H.C., Probst, C., Maier, A.: Enhanced crack segmentation (eCS): a reference algorithm for segmenting cracks in multicrystalline silicon solar cells. IEEE J. Photovolt. **9**(3), 752–758 (2019). <https://doi.org/10.1109/JPHOTOV.2019.2895808>
 90. Tsai, D.M., Wu, S.C., Li, W.C.: Defect detection of solar cells in electroluminescence images using Fourier image reconstruction. Sol. Energy Mater. Sol. Cells **99**, 250–262 (2012). <https://doi.org/10.1016/j.solmat.2011.12.007>
 91. Tsai, D.M., Wu, S.C., Chiu, W.Y.: Defect detection in solar modules using ICA basis images. IEEE Trans. Ind. Inform. **9**(1), 122–131 (2013). <https://doi.org/10.1109/TII.2012.2209663>
 92. Tsai, R.Y.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE J. Robot. Autom. **3**(4), 323–344 (1987). <https://doi.org/10.1109/JRA.1987.1087109>
 93. Tseng, D.C., Liu, Y.S., Chou, C.M.: Automatic finger interruption detection in electroluminescence images of multicrystalline solar cells. Math. Probl. Eng. **2015**, 1–12 (2015). <https://doi.org/10.1155/2015/879675>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Sergiu Deitsch received a masters degree in computer science from Nuremberg Institute of Technology in 2011. He is currently pursuing a Ph.D. degree with the Pattern Recognition Lab at the Friedrich-Alexander University Erlangen-Nürnberg. His research interests include various aspects of image processing, object detection and tracking, sensor fusion, and machine learning.

Claudia Buerhop-Lutz has been working as a scientist in the field of quality control of photovoltaics for more than 15 years. She is specialized in using imaging techniques, such as electroluminescence and thermography, for defect detection in solar modules.

Evgenii Sovetkin graduated in mathematics in 2013 and received his Ph.D. degree at RWTH University Aachen in 2020. From 2019, he is working in IEK-5, Institute of Photovoltaics at Forschungszentrum Jülich. His research interests include statistical modeling for high-dimensional data analysis and statistical modeling for physics.

Ansgar Steland received M.Sc. and Ph.D. degrees in mathematics from the Georg-August University of Göttingen, Germany. He held positions as an assistant at the Technical University of Berlin, Berlin, Germany, as a consultant in industry, as a postdoc at the European University Viadrina of Frankfurt/Oder, Germany, and as a lecturer at the Faculty of Mathematics at the Ruhr-University Bochum. Since 2006, he is Full Professor at RWTH Aachen University, Germany, where he holds the Chair of Stochastics at the Institute of Statistics. He is an Elected Member of the International Statistical Institute (ISI), acts as the speaker of the Society for Reliability, Quality and Safety, and chairs the Section on Statistics in Natural Sciences and Technology of the German Statistical Society (DStatG). He organized and co-organized several international workshops and invited sessions, edited several proceedings and special journal issues, and serves the community as associate editor. His current research interests are in nonparametric statistics and high-dimensional inference, statistical and machine learning, neural networks, change-point detection and sequential analysis, image processing, and applications in photovoltaics, quality control and econometrics.

Andreas Maier received his Ph.D. in 2009. From 2005 to 2009, he was working at the Pattern Recognition Lab at the Computer Science Department of the University of Erlangen-Nuremberg. His major research subject was medical signal processing in speech data. From 2009 to 2010, he started working on flat-panel C-arm CT as post-doctoral fellow at the Radiological Sciences Laboratory in the Department of Radiology at the Stanford University. From 2011 to 2012, he joined Siemens Healthcare as innovation project manager and was responsible for reconstruction topics in the Angiography and X-ray Business Unit. In 2012, he returned the University of Erlangen-Nuremberg as head of the Medical Reconstruction Group at the Pattern Recognition Lab. In 2015, he became professor and head of the Pattern Recognition Lab.

Florian Gallwitz graduated in computer science in 1994 and received his Ph.D. degree in computer science at the Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany, in 2001. From 2000 to 2009, he worked in the speech technology industry, as a co-founder and head of R&D of the Sympalog Voice Solutions GmbH. In 2010, he was appointed as a Professor for computer science and media at the Nuremberg Institute of Technology. His research interests include pattern recognition, speech and image processing, and deep learning.

Christian Riess received the Ph.D. degree in computer science from the Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany, in 2012. From 2013 to 2015, he was a Postdoc at the Radiological Sciences Laboratory, Stanford University, Stanford, CA, USA. Since 2015, he is the head of the Phase-Contrast X-ray Group at the Pattern Recognition Laboratory at FAU. Since 2016, he is senior researcher and head of the Multimedia Security Group at the IT Infrastructures Lab at FAU, and he received his habilitation from FAU in 2020. He served on the IEEE Information Forensics and Security Technical Committee 2017-2019 and received the IEEE Signal Processing Award in 2017. His research interests include all aspects of image processing and imaging, particularly with applications in image and video forensics, X-ray phase contrast, and computer vision.