Automated Defect Detection and Localization in Photovoltaic Cells Using Semantic Segmentation of Electroluminescence Images

Joseph Fioresi, Dylan J. Colvin, Rafaela Frota, Rohit Gupta, Mengjie Li, Hubert P. Seigneur, Shruti Vyas, Sofia Oliveira, Mubarak Shah, *Life Fellow, IEEE*, and Kristopher O. Davis, *Member, IEEE*

Abstract—**In this article, we propose a deep learning based semantic segmentation model that identifies and segments defects in electroluminescence (EL) images of silicon photovoltaic (PV) cells. The proposed model can differentiate between cracks, contact interruptions, cell interconnect failures, and contact corrosion for both multicrystalline and monocrystalline silicon cells. Our model utilizes a segmentation Deeplabv3 model with a ResNet-50 backbone. It was trained on 17,064 EL images including 256 physically realistic simulated images of PV cells generated to deal with class imbalance. While performing semantic segmentation for five defect classes, this model achieves a weighted F1-score of 0.95, an unweighted F1-score of 0.69, a pixel-level global accuracy of 95.4%, and a mean intersection over union score of 57.3%. In addition, we introduce the UCF EL Defect dataset, a large-scale dataset consisting of 17,064 EL images, which will be publicly available for use by the PV and computer vision research communities.**

Index Terms—**Photovoltaics, electroluminescence imaging, defect classification, machine learning, semantic segmentation, deep learning.**

I. INTRODUCTION

PHOTOVOLTAICS (PV) has grown to be a multibillion dollar renewable energy industry in large part due to manufacturing line quality control. Performance measurements (current–voltage or *I–V*) and electroluminescence (EL) imaging are the most common characterization techniques for analyzing PV cell and module performance. *I–V* curves measure the electrical performance (i.e., efficiency) of both PV cells and modules. Photoluminescence and EL imaging can both provide a spatial representation of the local excess carrier density, and therefore local voltage, of wafers [1]–[3], cells [4]–[15], and modules [16]–[18]. EL imaging is heavily relied upon in PV module manufacturing, in particular, and it is used extensively to study the reliability and durability of PV modules when exposed to environmental stressors. However, EL image inspection is time-consuming and therefore costly task that requires experienced people to visually analyze thousands of images for accurate quality control. Machine learning provides an automated, cost-effective, and time-efficient solution.

The need for an automated defect detection model has certainly been recognized by the field. A sizeable amount of work has been done to create models to automatically find defects in PV modules using EL images, but the majority of those models either have a limited number of defect classes that can be distinguished and/or they do not provide pixel-level localization of the defects.

In 2012, Tsai *et al.* [19] used Fourier image reconstruction on EL images of multicrystalline c-Si cells to segment out small cracks, breaks, and finger interruptions. This resulted in a successful model that segmented these defects in 0.29 s per image without the use of machine learning. After this, deep learning methods started gaining traction with the development of more robust image processing models. However, Chen *et al.* [20] recognized that deep learning takes a lot of training samples and labeled data, so they used intense computation instead of a deep learning network. They were able to segment cracks in

---

Manuscript received June 15, 2021; revised November 19, 2021; accepted November 22, 2021. Date of publication December 14, 2021; date of current version December 23, 2021. This work was supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy under the Solar Energy Technologies Office under Agreement DE-EE0008155. The work of Joe Fioresi was supported under Grant NSF CNS-1757858. *(Corresponding authors: Joe Fioresi; Kristopher O. Davis.)*

Joseph Fioresi is with the Department of Computer Science, University of Central Florida, Orlando, FL 32816 USA, with Resilient, Intelligent, and Sustainable Energy Systems Cluster, University of Central Florida, Orlando, FL 32816 USA, and also with the Center for Research in Computer Vision, University of Central Florida, Orlando, FL 32816 USA (e-mail: joefioresi718@gmail.com).

Dylan J. Colvin and Mengjie Li are with Resilient, Intelligent, and Sustainable Energy Systems Cluster, University of Central Florida, Orlando, FL 32816 USA, with the Department of Materials Science and Engineering, University of Central Florida, Orlando, FL 32816 USA, and also with FSEC Energy Research Center, Cocoa, FL 32922 USA (e-mail: dylanjc43@knights.ucf.edu; mengjie.li@ucf.edu).

Rafaela Frota is with Resilient, Intelligent, and Sustainable Energy Systems Cluster, University of Central Florida, Orlando, FL 32816 USA, and also with the College of Optics and Photonics, University of Central Florida, Orlando, FL 32816 USA (e-mail: rafaelafrota@knights.ucf.edu).

Rohit Gupta, Shruti Vyas, and Mubarak Shah are with the Department of Computer Science, University of Central Florida, Orlando, FL 32816 USA, and also with the Center for Research in Computer Vision, University of Central Florida, Orlando, FL 32816 USA (e-mail: rohitg@knights.ucf.edu; shruti@crcv.ucf.edu; shah@crcv.ucf.edu).

Hubert P. Seigneur is with FSEC Energy Research Center, Cocoa, FL 32922 USA (e-mail: seigneur@ucf.edu).

Sofia Oliveira is with Resilient, Intelligent, and Sustainable Energy Systems Cluster, University of Central Florida, Orlando, FL 32816 USA (e-mail: soliveira@knights.ucf.edu).

Kristopher O. Davis is with Resilient, Intelligent, and Sustainable Energy Systems Cluster, University of Central Florida, Orlando, FL 32816 USA, with the Department of Materials Science and Engineering, University of Central Florida, Orlando, FL 32816 USA, with FSEC Energy Research Center, Cocoa, FL 32922 USA, and also with the College of Optics and Photonics, University of Central Florida, Orlando, FL 32816 USA (e-mail: kristopher.davis@ucf.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JPHOTOV.2021.3131059.

Digital Object Identifier 10.1109/JPHOTOV.2021.3131059

multicrystalline cells with 0.949 F1-score and a fast calculation time of 0.053 s.

Tseng *et al.* [21] used a spectral clustering algorithm to both identify and classify contact fingers as interrupted or noninterrupted. Their method finds almost all interrupted fingers with an accuracy rate of over 99%; an impressive result, but limited to only one specific type of PV cell manufacturing defect.

Deitsch *et al.* [22] proposed the use of a convolutional neural network (CNN) in this application, performing binary classification on whether or not a defect was present in each cell with 88.42% accuracy. This group is also responsible for the public release of a benchmark labeled dataset of 2624 EL images of c-Si solar cells segmented from the EL module images [22]–[24]. Akram *et al.* [25] used the above-mentioned dataset in their work. They show the viability of a light CNN for binary classification, achieving high accuracy and low inference times. Their light structure beat out the original deep CNN, getting 93.02% classification accuracy in only 8.07 ms per image.

Karimi *et al.* [26] also used a CNN on their dataset to classify cells into five defect categories (good, corroded, cracked, between busbar darkening, and cell edge darkening) to achieve 98.24% accuracy. The same group also tried an end-to-end defect detection pipeline later on. They preprocessed modules into individual cell images for classification into three categories (good, cracked, and corroded). After ending the pipeline with a CNN, SVM, and a random forest model, the CNN outperformed both of the other machine learning models with a 99.71% classification accuracy [27].

Most of the deep learning models in this field had an issue with small datasets. Tang *et al.* [28] tried to counter this with advanced data augmentation using a generative adversarial network (GAN) and got positive results. Using a base eight-layer CNN without any form of data augmentation saw less than 50% classification accuracy for each defect category. With their proposed data augmentation methods, each category saw over 80% accuracy, the lowest being finger interruptions at 81%.

Most work has dealt with image-level defect recognition and classification. This classification is useful, but looking at the pixel level can provide more granular information. Otamendi *et al.* [29] created an end-to-end anomaly detection pipeline that ended in binary pixel-level segmentation. This segmentation identifies cracks, microcracks, dead spots, weak areas, and weak cells; however, it does not separate these categories automatically, whereas our work separates defect segmentations into unique categories. Mayr *et al.* [30] performed EL crack segmentation using image-level annotations and normalized $\mathrm{L_p}$ norms for the best crack segmentation F1-score of 0.83, which is close to our crack category F1-score of 0.81. Zhao *et al.* [31] tried bounding box segmentation for 14 defects useful in production line analysis, achieving 70.2% $\mathrm{mAP_{50}}$ for the whole validation set. This work succeeded in identifying and localizing those defects although our semantic segmentation method is more exact in the localization.

In summary, while many works in the field have shown success in automated defect detection in PV modules from EL images, this article demonstrates a sophisticated deep learning model that individually classifies and localizes four different combined defect categories using semantic segmentation with the possibility of moving to nine defect classes in the near future. This work lays out a new state of the art for multicategory pixel-level defect segmentation of EL images of PV modules. This is accomplished in large part by constructing the UCF EL Defect dataset, a dataset large enough to overcome the common low-data problem and the supplementation of the dataset using simulated images generated using an established finite element model grounded in physics. The dataset of 17 064 full annotated EL images will be made public, along with the model itself.

II. DATASET

Our UCF-EL-Defect dataset currently contains 17 064 annotated cell images [35]. This number is much higher compared to similar existing datasets available for this task. As shown in Table I, this dataset will be publicly available for the research community. Before this, there were only two notable public datasets, namely the ELPV dataset [23] and Sovetkin *et al.* [34]. Our dataset has more images as well as defect classes. Pixel-level annotations are available to enable semantic segmentation of nine defect classes.

Defects are categorized into nine classes for annotation. These classes are distinguished by their unique EL profile and location on the cell surface. The nine classes are illustrated in Fig. 1, which are as follows: closed cracks; resistive cracks, isolated cracks, front grid interruptions, grid interruptions near the busbar, disconnected interconnects, highly resistive interconnects, and contact corrosion.

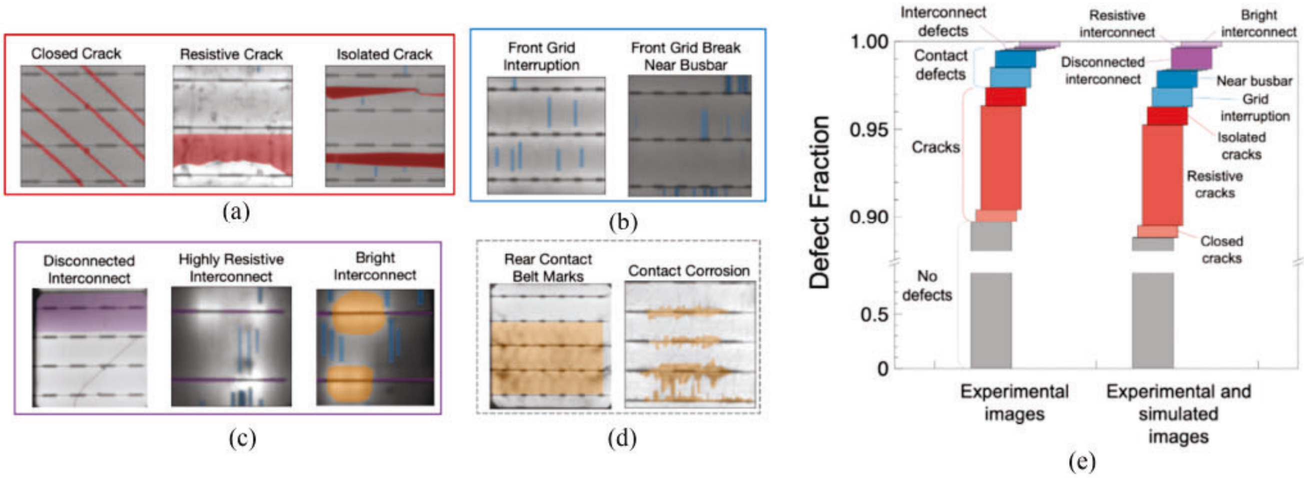
Cracks and cell fracture have been well studied by the PV research community and can take many forms [17], [36]–[45]. Closed cracks appear as black, narrow lines, whereas annotations for resistive cracks include the closed crack lines along with darkened regions extending from the crack lines. Isolated cracks refer to regions that are clearly individual sections of bulk material that have been split from the main cell by cracks. These cracks are visualized in Fig. 1(a) in their respective order. Resistive and isolated cracks are most easily distinguished by EL profile. Resistive cracks have a gradient in intensity, whereas isolated have a clearly uniform intensity across the surface.

Front grid interruptions in Fig. 1(b) appear as darkened rectangles appearing along finger lines [11]. These are normally the result of a screen-printing issue and can therefore happen at various locations on the cell. They can also be caused by cracks leading to a discontinuity in the metal contact. Often, these interruptions can occur near the busbar, and this case is labeled separately because it is often caused by cracks forming along the busbar resulting from high stress during stringing and tabbing and/or excessive mechanical loads.

Disconnected interconnections, such as the first cell in Fig. 1(c), appear as long dark rectangles that extend along the entire length of the busbar [46], [47]. Disconnections on interconnects along edge busbars appear darker than those in the center. Highly resistive interconnects are brighter either along

TABLE I
COMPARISON OF OUR EL IMAGE DATASET WITH VARIOUS DATASETS USED IN EL IMAGE BASED DEFECT DETECTION OR CLASSIFICATION STUDIES

Dataset (access)	# of Images	Defect classes	Annotations	PV Cell Technology
Karimi et al. [27] (No)	5,400 cell images, 3,550 annotated	Three: cracked, corroded, and good	Single label classification	monocrystalline
Su et al. [32], [33] (No)	3,629 annotated images	Four: crack, finger interruption, and black core	Classification and bounding box with IoU values	multicrystalline
ELPV-dataset [23] (Yes)	2,426 annotated images	Two: Defective and functional	Defect probability and cell type	monocrystalline/ multicrystalline
Tang et al. [28] (No)	1,800 images, 450 each defect	Four: micro crack, finger interruption, break, and defect free	Classification	monocrystalline/ multicrystalline
Chen et al. [20] (No)	10,000 images	One: crack	Segmentation	multicrystalline
Sovetkin et al. [34] (Yes)	156 annotated module images, 6,000 unlabelled	Two: shunts and droplets	Semantic segmentation	CIGS
UCF EL Defect Dataset (this work) (Yes)	17,064 annotated cell images	Nine defect classes	Semantic segmentation	monocrystalline/ multicrystalline



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

Fig. 1. Examples of each defect category annotated for. Consistent colors are used to indicate the grouping the defect belongs to. Class distributions with and without simulated images: each bar shift represents a different defect, with the grouped colors/labels representing the way defects are grouped. The purple interconnect class distribution is greatly increased with simulated images.

and perpendicular to the length of the busbar, or bright along part of the busbar and dark across the remaining length [17], [48]. An example is shown in the middle interconnect defect cell in Fig. 1(c). When the entire busbar is bright, the bright EL response extends a short distance perpendicular to the busbar, but extends farther when it is only part of the busbar. Several of the highly resistive interconnects in our dataset do not affect the entire length of the busbar. This gives shorter (less distance across the busbar) but wider (longer perpendicular distance) bright regions. The bright spots distract the segmentation from appropriately labeling the defect, so the bright spots were listed as their nondefective class so the model would better distinguish the two interconnect defects. An example of this type of bright region is also shown in Fig. 1(c).

Some more rare contact defects were also labeled. Rear contact belt marks are due to nonuniform firing of the rear Al paste. The darker regions observed are normally the location where the moving belt of the fast firing furnace contacts the paste. These were not considered in the work because they are not really critical defects. Contact corrosion, shown as the second cell in Fig. 1(d), has a characteristic darkening pattern starting at the busbar and extending perpendicularly [49]–[51]. The pattern extends out farther from the center of the busbar than near the edges. These defects only emerge after exposure in the field or, more likely, after accelerated aging in damp heat conditions.

A. Class Imbalance

The full dataset contains starts with 16 808 unique cells parsed from 368 crystalline silicon (c-Si) PV modules. The high number of images combined with the relatively low size of each defect results in a highly imbalanced dataset. Defect class distributions were very low compared to total pixel counts. This creates a challenge for the model to learn defects without a significant amount of instances in our dataset. Defects were merged into four categories to boost class distributions. Closed cracks, resistive cracks, and isolated cracks are combined into a crack category. Front grid interruptions and interruptions near the busbar make up the contact category. Highly resistive interconnects get their category due to the lack of annotated disconnected interconnects. Contact corrosion also gets a unique category due to plenty of examples. Disconnected interconnects are currently ignored with plans to include them upon further annotations.

Undersampling is a useful method for reducing the number of examples in the overwhelming class. We ignore images without defects to help balance classes. With undersampling, the total amount of defective pixels of all categories combined is only 10%. Particularly, interconnect-related defects made up less than 0.2% of pixels in the dataset. Contact interruptions only made up just over 2%. Due to this, 256 physically realistic simulated defective cell images were created. These simulated images focused on both interconnect and contact defects, raising interconnect defect class distribution to 1.2% while maintaining roughly the same percentage of contact defects. The addition of simulated images results in a final dataset of 17 064 EL images.

B. Annotations

Annotations were done using the VGG Image Annotator, [52], [53] an open-source annotation software created by a group at the University of Oxford, Oxford, U.K. The defects on the EL images were manually outlined pixel by pixel, providing a more accurate annotation of different defects. The annotations are saved in a comprehensive .csv file with a filename, file size, region count (per cell), region count index, defect shape/location, and defect type. This file will be made publicly available with the dataset. Cell damage ranged from no damage to severe, with severely damaged cells displaying multiple classes of defects in the same image. Defects were annotated with little overlap between the defects. All defects were annotated into separate categories, and the groupings occur when processing the annotations into ground truth images. A custom program is used before training to convert annotations into ground truth segmentation images. Of the total images annotated, the number of images with no defects is 4 114, whereas those with at least one contact, crack, and interconnect defect are 10 038, 4707, and 476, respectively.

C. Generating Simulated Images

A physics-based solar cell device simulation software, Griddler-Pro [52], was used to generate realistic simulated images. We first perform a calibration step. The calibration step is done by reproducing the experimental images. Images with heavy grid interruptions and disconnected interconnects were chosen. After the calibration step, variation on the particular defect classes (such as location and distribution) was introduced to generate more simulated images. Diagrams and more specific descriptions of this process may be found in [53].

D. EL Imaging

EL imaging operates under the principle of radiative recombination. A power supply injects carriers into the PV module and recombine radiatively to emit photons corresponding to the bandgap energy of the luminescencing material. Due to the bandgap of Si (1.12 eV), the emitted light is in the near IR range for (1180 nm). Sensors register the photons as signals and convert them into an image that can be used for qualitative inspection or quantitative analysis. The two cameras used in this work are a 16 MP modified DSLR CMOS with a 950-nm long-pass filter and an 8 MP cooled Si charge-coupled device (CCD) with an 850-nm long-pass filter. The cooled CCD improves SNR and enhances feature visibility by minimizing the thermal noise within the sensor.

Different module types can exhibit different characteristic patterns in EL images, particularly due to wafer crystallinities. The module types represented in our dataset are multicrystalline and monocrystalline aluminum back surface field cells, as well as monocrystalline passivated emitter and rear cells. The interconnection scheme represented is conventional ribbon tabbing with three to five busbars.

III. METHODS

A. Network Architecture

The model used is a pretrained DeepLabv3 [54] model with a ResNet50 [55] backbone. Based on the success in previous semantic segmentation works [56], we experimented with DeepLabv3 with ResNet50 and ResNet101 backbone, and FCN with ResNet50 and ResNet101 backbone. DeepLabv3_ResNet50 was selected as our final model as it showed the best performance. DeepLabv3 uses atrous convolution and outperforms its previous versions. The model was pretrained on the Microsoft COCO dataset [57] and modified as per our dataset. The pretrained model had 21 classes, whereas this work has five (four defects plus the interconnect bright spot), so in the final model, modifications have been made to accommodate this difference. Images of at least 224 × 224 pixels are used as inputs. As seen in the workflow in Fig. 2, the images are input to Deeplabv3. Being an encoder–decoder type network, it creates a feature representation of the input image. The features contain the relevant information for classification and segmentation of the defects. These features are upsampled through our ResNet50 based decoder into a semantic segmentation image of the same size as the input image. The pixels are individually assigned a label corresponding to a defect class, chosen by the final softmax layer.

B. Image Preprocessing

We start by accepting the image of a cell along with its corresponding annotation. These are both converted to PyTorch tensors for efficient use. The images are split 80–20 into training and validation/testing sets. The simulated images are only included in the training set as they do not reflect the performance on real-world data. In our dataset, the images are grayscale and of varying sizes, but generally around 300 × 300 pixels. Every image is normalized, randomly resized, and randomly center cropped for consistent input into the model (224 × 224). Even though we are working with a large dataset, we use certain data augmentation techniques to further improve the performance. We add three random transforms and one blur transform to every training image to help the model generalize. The training images have independent 50% chances to be horizontally flipped, vertically flipped, and rotated 90°. A Gaussian blur with a 7 × 7 kernel and a standard deviation of 2 is applied to the training set as well. By training on all these different preprocessing transforms, we expect the model to be robust to minor changes in viewpoint and quality of input images.

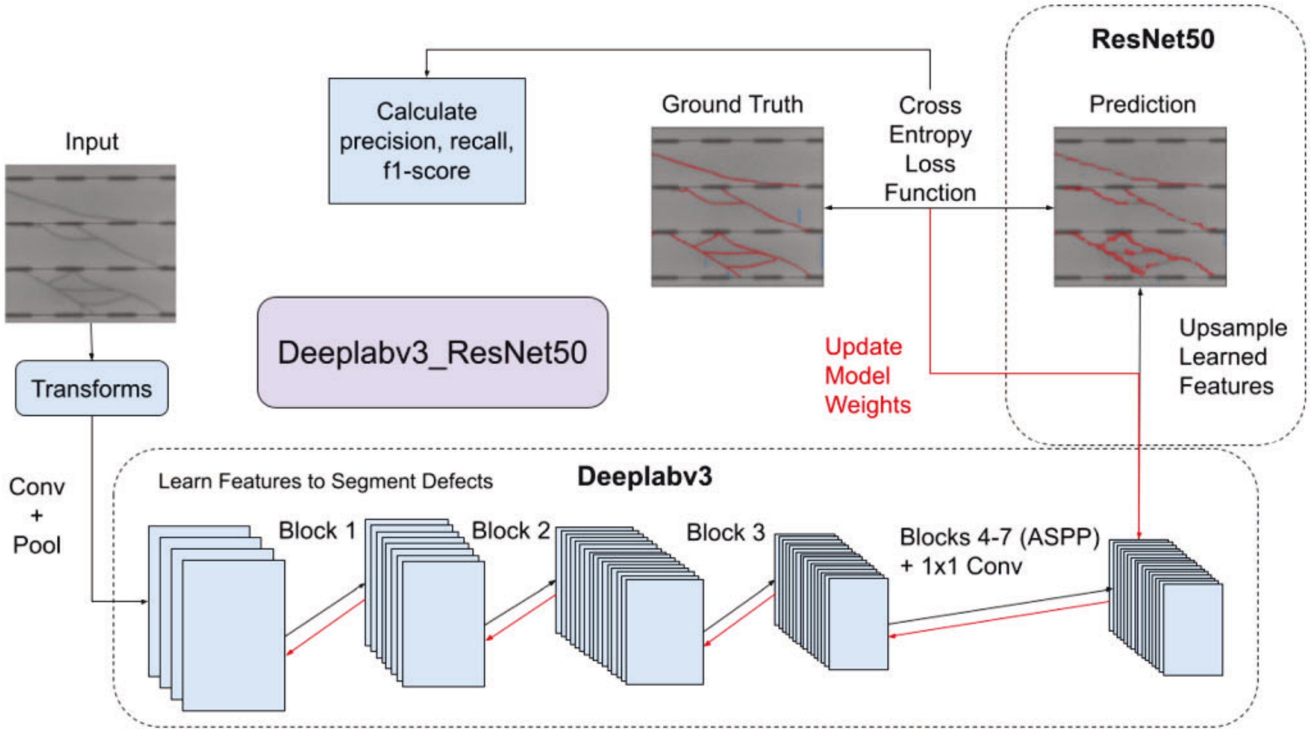


Fig. 2. Visual of the model workflow: Start with an input cell image, apply transforms, then first convolution and pooling layers. The image propagates forward through the Deeplabv3 model blocks to learn features for identifying and localizing the defects. These resulting features are upsampled using the ResNet 50 backbone to produce a segmented defect prediction image. A loss function compares this to the ground truth annotation. The result of this is simultaneously used to calculate test metrics and to update the learned model weights by backpropagating through the model blocks (shown in red).

C. Model Training

After the preprocessing, for training, we input the transformed images with their annotation as the ground truth. A visual workflow of this is shown in Fig. 2. The model segments each image, labeling every pixel as defect-free, a crack defect, a contact defect, a highly resistive interconnect, or contact corrosion. This segmentation is then compared to the preannotated ground truth with a cross-entropy loss function. The model weights at each layer are updated with the goal of reducing that loss toward a minimum. After each epoch, we validate the model on the validation/testing set to monitor the model training. The hyperparameters were optimized for best performance using a hit and trial approach. For every input image, we get a segmentation map of the same size. Zeroes are assigned to pixels with no defect and numbers at pixels corresponding to specific defect categories. This can be visualized by holding a transparent defect map over the original image, with proper alignment, to highlight the defects in their appropriate locations.

D. Model Adjustments

The types of cells in this dataset have a relatively high degree in variation, notably seen in monocrystalline versus multicrystalline Si cells. The monocrystalline cells have very little background noise and are visually clear. Multicrystalline cells have noisy backgrounds that sometimes mask defects. In the multicrystalline cells especially, the model is not very confident about small or dim defects. The model ended up being much more confident about predicting pixels as nondefective than defective. Originally, the final layer of the neural network was as a simple softmax, taking the highest label probability and assigning the label from there. Smaller/dimmer defects were often ignored in favor of nondefective due to low confidence. To pick up on these defects more, we make a major adjustment to the model by removing the final softmax layer. This allows us to utilize the raw predicted probability of a defect class existing at each pixel. The label probability for nondefect is compared to a custom threshold. If that probability is above the threshold, the pixel is labeled as nondefective. Otherwise, a softmax is applied to the rest of the defects, and the pixel is labeled as the highest defect probability. A high threshold is chosen to ensure that the model is very confident and there is no defect. This method saw an increase in recall of defects as using an appropriate threshold allowed the model to find less prominent defects.

E. Evaluation

The main metrics discussed in this work are precision, recall, and F1-score. These all are related to true and false positives (TP, FP) and negatives (TN, FN). Global pixel accuracy and IoU scores are also mentioned

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{TP + FN}$$

$$F1 - \text{score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{precision} + \text{recall}}$$

TABLE II
RESULTS COMPARISON TRAINING WITH AND WITHOUT SIMULATED IMAGES

	Without Simulated			With Simulated		
# of Images	16,808			17,064		
	Result Metrics					
	prec	recall	f1	prec	recall	f1
Crack	0.81	0.80	0.80	0.81	0.80	**0.81**
Contact	0.68	0.46	0.55	0.66	0.51	**0.58**
Interconnect	0.14	0.34	0.20	0.26	0.44	**0.33**
Corrosion	0.73	0.79	0.76	0.69	0.89	**0.78**

Bold f1-scores indicate the best score between the two experiments.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}}.$$

A cross-entropy loss function used during training to compare ground truth to predicted images is

$$\text{loss}(x, \text{class}) = -x[\text{class}] + \log\left(\sum_j \exp(x[j])\right).$$

The model works to minimize the output of the above equation. In short, the closer the prediction is to the ground truth, the lower the loss will be. To specify penalty for missing certain classes, the right-hand side may be multiplied by a preset weight for each category if specified.

IV. RESULTS

A. Model Performance

The proposed model attains per-class F1 scores of 0.98, 0.81, 0.58, 0.33, and 0.78 for no-defect, crack, contact, interconnect, and corrosion defects, respectively (see Table II). These were achieved using Deeplabv3_ResNet50 modified to our dataset while using the following group of important hyperparameters: learning rate = 0.005, threshold = 0.54, and using auxiliary loss. Scikit-learn classification reports are recorded for each testing epoch and show important results for performance comparison.

B. Impact of Simulated Images

Due to severe class imbalance, our model results may be improved with the addition of artificially generated image samples using a GAN. This work forgoes this common solution as the images in the dataset are tied to physical calculations that cannot be easily generated. We instead maintain the physically realistic calculations by carefully simulating cells using Griddler-Pro [52].

After establishing the baseline, 256 simulated images generated by physics-based modeling were added to the database, resulting in total 17 064 images. The ratio of simulated images is small, but they had a notable effect on the interconnect and contact category distributions. Within the simulated images, the distribution of pixels was 43.5% nondefective, 20% grid interruption, and 36.5% interconnect defective. The results with and without the simulated images are shown below in Table II. Even with a small batch of simulated images like this added to the database, improvement on the class-imbalance and the machine learning model accuracy was observed. The deep learning model accuracy was also improved. The F1 scores of contact and interconnect categories were boosted while maintaining similar scores in the crack and corrosion categories.

C. Qualitative Analysis

Fig. 3 shows six examples of how the model performs. The first image is the annotated ground-truth, whereas the second image is the model segmentation prediction output. Fig. 3(a) shows the model successfully segmenting corrosion, and high results in this category show that it is able to successfully find most corrosion. Corrosion can appear visually similar to the noisy multicrystalline background, but the results show that this is not an issue. Fig. 3(b) shows a simple example of all contact interruptions being found on a monocrystalline cell. Fig. 3(c) shows a contrasted example of highly resistive interconnects, which the model is able to segment without the help of the bright spot class. It even finds the contact interruptions hidden in the dark. However, it is important to note that the interconnect annotation went through the bright spot in this case. Even with all three different types of cracks merged into a single crack category, the model is able to find each and highlight the defective area properly, even within the same image, as demonstrated in Fig. 3(d). Fig. 3(e) shows that the model is capable of distinguishing between similar-looking defects (resistance bars stemming from a crack versus contact interruptions near the solder pad). The final example [see Fig. 3(f)] shows another example of a cell with many defects where the model is able to accurately identify and localize all of them.

D. Model Shortcomings

The major shortcoming of this model is its dependence on supervision. The only source of learning is from annotated images. This makes it difficult to generalize outside of the experimental set. Like other supervised methods, it does well dealing with cell and defect types it was trained on, however, it is not usable with new inputs or unseen classes. Some of the cell types have very few samples in training data and thus are seen relatively few times in the training. The model is able to detect the defect in such cases, however, their segmentation is flawed. Fig. 4(c) is an example of cracks that might be missed in these cases. Overall, the model does well at generalizing across cell types, but loses out on some accuracy this way. Keeping the input images a consistent type may result in higher accuracies, however, lowering the generalization abilities. Also, in case of very close defects, the model tends to lump the segmentations together. The predictions are rounded off and not very distinct, but ultimately still accurately find and localize the defects.

This model also misses some faint contact grid interruptions. Combining faint defects with a noisy multicrystalline background distracts the model, causing it to miss contact interruptions [see Fig. 4(a)]. The grid interruptions have a somewhat large degree of intraclass variation, causing the model to lose confidence in smaller instances. Sometimes, the model even overwrites a contact annotation with a crack annotation [see

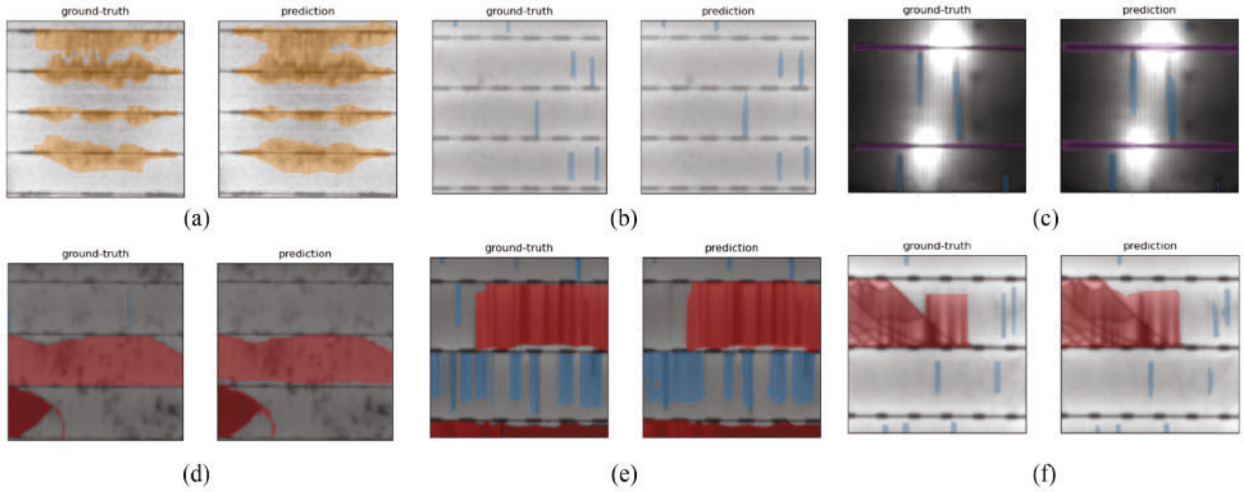


Fig. 3. Example visualized output of the model, red = crack, blue = contact, purple = interconnect, orange = corrosion. (a) Solid corrosion segmentation. (b) Faint contact defects in a monocrystalline cell. (c) Highly resistive interconnect without bright spot. (d) All three types of cracks with a faint resistive boundary. (e) Spanning resistive crack mixed with contact interruptions. (f) Crack conglomeration + faint contact interruptions.

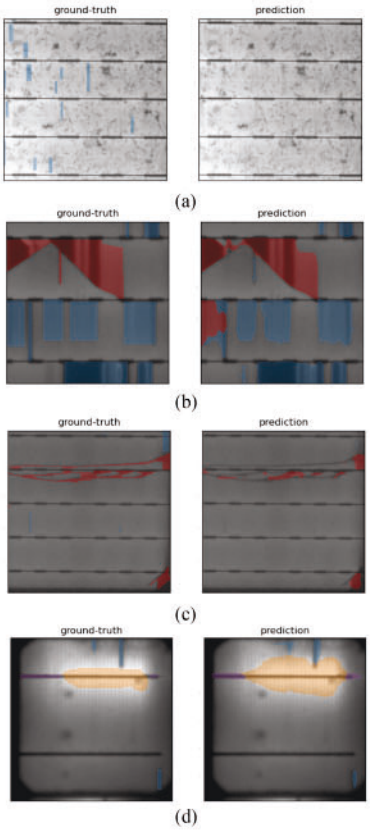


Fig. 4. Examples of shortcomings. (a) Failed finding faint contact interruptions in a multicrystalline cell. (b) Overwrite blue contact annotation with red crack annotation. (c) Visually obvious crack being missed. (d) Interconnect annotation going through bright spot.

Fig. 4(b)]. We addressed this by weighting the contact class higher in the loss function, but this ended up causing the model to predict too many contact interruptions. We believe that this issue can be addressed with more fine-tuning.

The partial high resistance interconnects posed a particularly difficult challenge for our model. The visual indication of the defect is a bright spot. The actual partial high resistance defect occurs in exactly the parts of the busbar not covered by the bright spot. In some cases, the model would annotate the busbar through the bright spot [see Fig. 3(c)]. However, as our model is meant to localize defects on top of identifying them, we wanted to cut out this incorrect labeling. As a solution, we added a bright spot "defect" class, even though it is not its own defect. The goal was to get the model to recognize the bright spots and to pair them with resistive interconnects. As a result, the bright spots were found by the model, but it did not seem to always pair it with a highly resistive interconnect failure. In fact, the bright spot prediction sometimes took over the busbar, overwriting potential resistive interconnect defects, as shown in Fig. 4(d). This actually lowered the prediction scores of the interconnect category from the best F1-score of 0.44 without bright spots to the best F1-score of 0.20. As a result, the best models are trained ignoring the bright spot. However, the annotations remain and may be useful in future works.

V. DISCUSSION

Each image takes an average of 18.1 ms to process through the model. This is slower than the lightweight classification model in [25], which achieves 8.07 ms per image. However, it is faster than existing notable computational (non-ML) segmentation methods (290 ms, 53 ms) [19], [20].

This dataset contains a high degree of module variation. There are visual differences between different modules, such as the number of busbars, different manufacturers, and different cell technologies and wafer crystal structures (i.e., mono and multi). A strength of this dataset is the variation, which ensures that the model learns to generalize defects despite differences in modules. However, the generalization contributes to worse results than if the model were trained on a consistent dataset, but this sacrifice improves the model's overall robustness.

Our model uses a cross-entropy loss function to update the network weights, with each defect class having an equal weight when calculating loss. We tried assigning higher weights to the smaller defect categories to penalize the model more severely for missing the defects. We calculated these weights by taking the inverse of their distribution percentages. This resulted in very low weights for no-defect pixels and very high weights for defective pixels. This method negligibly improved the model.

Another loss function modification we tried was scaling loss weight to instance size. The cross-entropy function is not reduced, and we instead multiply the nonreduced loss pixels by their calculated instance weights. These weights are calculated by taking (1—inverse of instance size). These are averaged out into a final loss value. With this customized cross-entropy loss, small instances of defects had higher loss weights than larger instances. The dataset does not take care to separate defects by instance and some instances are lumped together although there are small defects that could potentially benefit from a higher loss weight. This method negligibly improved the model although modifying the weight calculation may improve the model further.

VI. CONCLUSION

In this article, we present a semantic segmentation deep learning model that identifies and localizes defects in EL images. To the best of our knowledge, this is the first work to segment these specific defect categories at a pixel level. The model is able to segment defects across numerous different crystalline Si cell technologies. We also introduce the 17 064 EL image UCF EL Defect dataset, fully annotated for nine defect categories. Using this dataset, the current model is able to identify and localize nondefective area, cracks, contact defects, corrosion, and interconnect defects out with an average of 0.69 F1-score and a pixel-level global accuracy of 95.4%. There is potential to use our model for industry EL inspection to extract more information and reduce resources and time necessary for inspection.

Our efforts focus on simultaneously increasing the current defect category scores and adding in more defect categories at similar accuracy values. We also plan on further splitting up the current groupings into individual categories as it allows for a more accurate interpretation of cell defects. Future work will focus on relating the defect area fractions to other PV module data, particularly *I–V* curves, in the hopes of distributing power loss to each individual defect type.

ACKNOWLEDGMENT

The authors would like to thank Clark Rodriquez, Cristian Pearson, George Hutchinson, Makayla Dominger, Nancy Zheng, Nicholas Moser-Mancewicz, Nicole Karam Pannaci, and Sofia Oliveira for their work annotating the dataset. They also would like to thank Andrew Gabor of BrightSpot Automation, Greg Horner of Tau Science, Eric Schneller of Silfab Solar, and William Hobbs of Southern Company for their helpful feedback in selecting the appropriate defect categories.

REFERENCES

[1] T. Trupke, J. Nyhus, and J. Haunschild, "Luminescence imaging for inline characterisation in silicon photovoltaics," *Physica Status Solidi Rapid Res. Lett.*, vol. 5, pp. 131–137, 2011.

[2] T. Niewelt *et al.*, "Interstitial oxygen imaging from thermal donor growth-a fast photoluminescence based method," *Sol. Energy Mater. Sol. Cells*, vol. 131, pp. 117–123, 2014.

[3] H. Seigneur *et al.*, "Manufacturing metrology for c-Si photovoltaic module reliability, and durability, Part i: Feedstock, crystallization and wafering," *Renewable Sustain. Energy Rev.*, vol. 59, pp. 84–106, 2016.

[4] T. Trupke, E. Pink, R. A. Bardos, and M. D. Abbott, "Spatially resolved series resistance of silicon solar cells obtained from luminescence imaging," *Appl. Phys. Lett.*, vol. 90, 2007, Art. no. 093506.

[5] D. Hinken, K. Ramspeck, K. Bothe, B. Fischer, and R. Brendel, "Series resistance imaging of solar cells by voltage dependent electroluminescence," *Appl. Phys. Lett.*, vol. 91, no. 18, Oct. 2007, Art. no. 182104. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.2804562

[6] H. Kampwerth, T. Trupke, J. W. Weber, and Y. Augarten, "Advanced luminescence based effective series resistance imaging of silicon solar cells," *Appl. Phys. Lett.*, vol. 93, no. 20, Nov. 2008, Art. no. 202102. [Online]. Available: https://aip.scitation.org/doi/abs/10.1063/1.2982588

[7] J. Haunschild, M. Glatthaar, M. Kasemann, S. Rein, and E. R. Weber, "Fast series resistance imaging for silicon solar cells using electroluminescence: Fast series resistance imaging for silicon solar cells using electroluminescence," *Physica Status Solidi Rapid Res. Lett.*, vol. 3, no. 7/8, pp. 227–229, Oct. 2009. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/pssr.200903175

[8] M. Glatthaar *et al.*, "Evaluating luminescence based voltage images of silicon solar cells," *J. Appl. Phys.*, vol. 108, 2010, Art. no. 014501.

[9] H. Höffler, H. Al-Mohtaseb, J. Haunschild, B. Michl, and M. Kasemann, "Voltage calibration of luminescence images of silicon solar cells," *J. Appl. Phys.*, vol. 115, 2014, Art. no. 034508.

[10] B. Hallam, B. Tjahjono, T. Trupke, and S. Wenham, "Photoluminescence imaging for determining the spatially resolved implied open circuit voltage of silicon solar cells," *J. Appl. Phys.*, vol. 115, 2014, Art. no. 044901.

[11] K. O. Davis *et al.*, "Manufacturing metrology for c-Si module reliability and durability Part II: Cell manufacturing," *Renewable Sustain. Energy Rev.*, vol. 59, pp. 225–252, 2016.

[12] M. J. Hossain *et al.*, "A comprehensive methodology to evaluate losses and process variations in silicon solar cell manufacturing," *IEEE J. Photovolt.*, vol. 9, no. 5, pp. 1350–1359, Sep. 2019.

[13] E. J. Schneller *et al.*, "Spatially resolved characterization of optical and recombination losses for different industrial silicon solar cell architectures," *Int. J. Mod. Phys. B*, Sep. 2020, Art. no. 2050204. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S0217979220502045

[14] K. Drabczyk, G. Kulesza-Matlak, A. Drygała, M. Szindler, and M. Lipiński, "Electroluminescence imaging for determining the influence of metallization parameters for solar cell metal contacts," *Sol. Energy*, vol. 126, pp. 14–21, Mar. 2016. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0038092X15007021

[15] G. Dost, H. Höffler, and J. M. Greulich, "Advanced series resistance imaging for silicon solar cells via electroluminescence," *Physica Status Solidi (a)*, vol. 218, no. 6, 2011, Art. no. 2000546. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pssa.202000546

[16] S. Guo, E. Schneller, K. O. Davis, and W. V. Schoenfeld, "Quantitative analysis of crystalline silicon wafer PV modules by electroluminescence imaging," in *Proc. IEEE 43rd Photovolt. Specialists Conf.*, 2016, pp. 3688–3692.

[17] E. J. Schneller *et al.*, "Manufacturing metrology for c-Si module reliability and durability Part III: Module manufacturing," *Renewable Sustain. Energy Rev.*, vol. 59, pp. 992–1016, 2016.

[18] D. B. Sulas, S. Johnston, and D. C. Jordan, "Comparison of photovoltaic module luminescence imaging techniques: Assessing the influence of lateral currents in high-efficiency device structures," *Sol. Energy Mater. Sol. Cells*, vol. 192, pp. 81–87, Apr. 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0927024818305907

[19] D.-M. Tsai, S.-C. Wu, and W.-C. Li, "Defect detection of solar cells in electroluminescence images using Fourier image reconstruction," *Sol. Energy Mater. Sol. Cells*, vol. 99, pp. 250–262, 2012.

[20] H. Chen, H. Zhao, D. Han, W. Liu, P. Chen, and K. Liu, "Structure-aware-based crack defect detection for multicrystalline solar cells," *Measurement*, vol. 151, 2020, Art. no. 107170.

[21] D.-C. Tseng, Y.-S. Liu, and C.-M. Chou, "Automatic finger interruption detection in electroluminescence images of multicrystalline solar cells," *Math. Problems Eng.*, vol. 2015, Oct. 2015, Art. no. 879675.

[22] S. Deitsch *et al.*, "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Sol. Energy*, vol. 185, pp. 455–468, 2019.

[23] C. Buerhop-Lutz, S. Deitsch, A. Maier, F. Gallwitz, and C. Brabec, "A benchmark for visual identification of defective solar cells in electroluminescence imagery," in *Proc. 35th Eur. PV Sol. Energy Conf. Exhib.*, 2018, pp. 1287–1289.

[24] S. Deitsch *et al.*, "Segmentation of photovoltaic module cells in uncalibrated electroluminescence images," *Mach. Vis. Appl.*, vol. 32, no. 4, Jul. 2021, Art. no. 84. *arXiv:1806.06530*. [Online]. Available: http://arxiv.org/abs/1806.06530.

[25] M. W. Akram *et al.*, "CNN based automatic detection of photovoltaic cell defects in electroluminescence images," *Energy*, vol. 189, 2019, Art. no. 116319.

[26] A. M. Karimi *et al.*, "Feature extraction, supervised and unsupervised machine learning classification of PV cell electroluminescence images," in *Proc. IEEE 7th World Conf. Photovolt. Energy Convers. (Joint Conf. 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC)*, 2018, pp. 0418–0424.

[27] A. M. Karimi *et al.*, "Automated pipeline for photovoltaic module electroluminescence image processing and degradation feature classification," *IEEE J. Photovolt.*, vol. 9, no. 5, pp. 1324–1335, Sep. 2019.

[28] W. Tang, Q. Yang, K. Xiong, and W. Yan, "Deep learning based automatic defect identification of photovoltaic module using electroluminescence images," *Sol. Energy*, vol. 201, pp. 453–460, 2020.

[29] U. Otamendi, I. Martinez, M. Quartulli, I. G. Olaizola, E. Viles, and W. Cambarau, "Segmentation of cell-level anomalies in electroluminescence images of photovoltaic modules," *Sol. Energy*, vol. 220, pp. 914–926, 2021.

[30] M. Mayr, M. Hoffmann, A. Maier, and V. Christlein, "Weakly supervised segmentation of cracks on solar cells using normalized lp norm," in *Proc. Int. Conf. Image Process.*, 2019, pp. 1885–1889.

[31] Y. Zhao, K. Zhan, Z. Wang, and W. Shen, "Deep learning-based automatic detection of multitype defects in photovoltaic modules and application in real production line," *Prog. Photovolt. Res. Appl.*, vol. 29, no. 4, pp. 471–484, 2021.

[32] B. Su *et al.*, "Deep learning-based solar-cell manufacturing defect detection with complementary attention network," *IEEE Trans. Ind. Informat.*, vol. 17, no. 6, pp. 4084–4095, Jun. 2021.

[33] B. Su, H. Chen, and Z. Zhou, "BAF-detector: An efficient CNN-based detector for photovoltaic solar cell defect detection," *IEEE Transact. Indust. Elect.*, 2021.

[34] E. Sovetkin, E. J. Achterberg, T. Weber, and B. E. Pieters, "Encoder-decoder semantic segmentation models for electroluminescence images of thin-film photovoltaic modules," *IEEE J. Photovolt.*, vol. 11, no. 2, pp. 444–452, Mar. 2021.

[35] "UCF-EL-Defect." 2021. [Online]. Available: https://github.com/ucf-photovoltaics/UCF-EL-Defect

[36] M. Köntges, I. Kunze, S. Kajari-Schröder, X. Breitenmoser, and B. Bjørneklett, "The risk of power loss in crystalline silicon based photovoltaic modules due to micro-cracks," *Sol. Energy Mater. Sol. Cells*, vol. 95, no. 4, pp. 1131–1137, Apr. 2011. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0927024810007129

[37] M. Dhimish, V. Holmes, B. Mehrdadi, and M. Dales, "The impact of cracks on photovoltaic power performance," *J. Sci. Adv. Mater. Devices*, vol. 2, no. 2, pp. 199–209, Jun. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2468217917300540

[38] E. J. Schneller *et al.*, "Evaluating solar cell fracture as a function of module mechanical loading conditions," in *Proc. IEEE 44th Photovolt. Specialists Conf.*, 2017, pp. 2897–2901.

[39] C. Buerhop *et al.*, "Evolution of cell cracks in PV-modules under field and laboratory conditions," *Prog. Photovolt., Res. Appl.*, vol. 26, no. 4, pp. 261–272, Apr. 2018. [Online]. Available: http://doi.wiley.com/10.1002/pip.2975

[40] F. Haase, J. Käsewieter, S. R. Nabavi, E. Jansen, R. Rolfes, and M. Köntges, "Fracture probability, crack patterns, and crack widths of multicrystalline silicon solar cells in PV modules during mechanical loading," *IEEE J. Photovolt.*, vol. 8, no. 6, pp. 1510–1524, Nov. 2018,

[41] E. J. Schneller, R. Frota, A. M. Gabor, J. Lincoln, H. Seigneur, and K. O. Davis, "Electroluminescence based metrics to assess the impact of cracks on photovoltaic module performance," in *Proc. IEEE 7th World Conf. Photovolt. Energy Convers. (Joint Conf. 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC)*, Jun. 2018, pp. 0455–0458.

[42] T. J. Silverman, M. Bliss, A. Abbas, T. Betts, M. Walls, and I. Repins, "Movement of cracked silicon solar cells during module temperature changes," in *Proc. IEEE 46th Photovolt. Specialists Conf.*, Jun. 2019, pp. 1517–1520. [Online]. Available: https://ieeexplore.ieee.org/document/8981150/

[43] C. M. Whitaker, B. G. Pierce, A. M. Karimi, R. H. French, and J. L. Braid, "PV cell cracks and impacts on electrical performance," in *Proc. 47th IEEE Photovolt. Specialists Conf.* , Jun. 2020, pp. 1417–1422. [Online]. Available: https://ieeexplore.ieee.org/document/9300374/

[44] M. Bdour, Z. Dalala, M. Al-Addous, A. Radaideh, and A. Al-Sadi, "A comprehensive evaluation on types of microcracks and possible effects on power degradation in photovoltaic solar panels," *Sustainability*, vol. 12, no. 16, Aug. 2020, Art. no. 6416. [Online]. Available: https://www.mdpi.com/2071-1050/12/16/6416

[45] L. Papargyri *et al.*, "Modelling and experimental investigations of microcracks in crystalline silicon photovoltaics: A review," *Renewable Energy*, vol. 145, pp. 2387–2408, Jan. 2020. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0960148119311589

[46] E. Annigoni *et al.*, "Quantifying and modeling the impact of interconnection failures on the electrical performance of crystalline silicon photovoltaic modules," *Prog. Photovolt., Res. Appl.*, vol. 27, no. 5, pp. 424–432, 2019. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/pip.3111

[47] D. J. Colvin, E. J. Schneller, and K. O. Davis, "Impact of interconnection failure on photovoltaic module performance," *Prog Photovolt., Res. Appl.*, vol. 29, no. 5, pp. 524–532, Mar. 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/10.1002/pip.3401

[48] R. Asadpour, D. B. Sulas-Kern, S. Johnston, J. Meydbray, and M. A. Alam, "Dark lock-in thermography identifies solder bond failure as the root cause of series resistance increase in fielded solar modules," *IEEE J. Photovolt.*, vol. 10, no. 5, pp. 1409–1416, Sep. 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9139421/

[49] J. L. Braid *et al.*, "EL and I-V correlation for degradation of PERC vs. Al-BSF commercial modules in accelerated exposures," in *Proc. IEEE 7th World Conf. Photovolt. Energy Convers. (Joint Conf. 45th IEEE PVSC, 28th PVSEC 34th EU PVSEC)*, 2018, pp. 1261–1266.

[50] A. J. Curran *et al.*, "Degradation of PERC and Al-BSF cells with UV cutoff and white variations of EVA and POE encapsulant," in *Proc. IEEE 48th Photovolt. Specialists Conf.*, Jun. 2021, pp. 1510–1516.

[51] N. Iqbal *et al.*, "Characterization of front contact degradation in monocrystalline and multicrystalline silicon photovoltaic modules following damp heat exposure," *Sol. Energy Mater. Sol. Cells*, vol. 235, Jan. 2022, Art. no. 111468. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0927024821005080

[52] A. Dutta, A. Gupta, and A. Zissermann, "VGG image annotator (VIA)," 2016. [Online]. Available: http://www.robots.ox.ac.uk/vgg/software/via/

[53] A. Dutta and A. Zisserman, "The VIA annotation software for images, audio and video," in *Proc. 27th ACM Int. Conf. Multim., ser. MM .19*. New York, NY, USA: ACM, 2019. [Online]. Available: https://doi.org/10.1145/3343031.3350535.

[54] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv:1706.05587*, 2017.

[55] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[56] S. Minaee, Y. Y. Boykov, F. Porikli, A. J. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2021.

[57] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *Proc. Comput. Vis.*, 2014, pp. 740–755.