

# WEAKLY SUPERVISED SEGMENTATION OF CRACKS ON SOLAR CELLS USING NORMALIZED $L_p$ NORM

Martin Mayr, Mathis Hoffmann, Andreas Maier, Vincent Christlein

Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nuremberg, Germany  
{martin.mayr, mathis.hoffmann, andreas.maier, vincent.christlein}@fau.de

## ABSTRACT

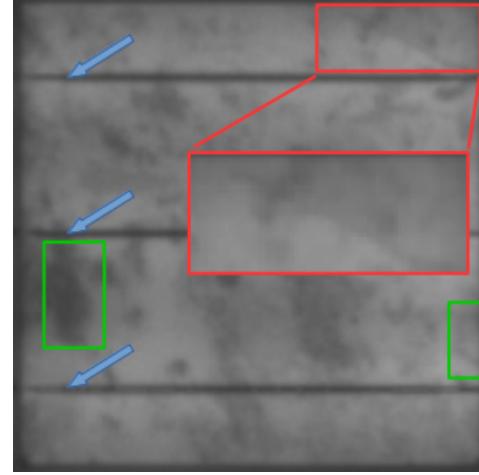
Photovoltaic is one of the most important renewable energy sources for dealing with world-wide steadily increasing energy consumption. This raises the demand for fast and scalable automatic quality management during production and operation. However, the detection and segmentation of cracks on electroluminescence (EL) images of mono- or polycrystalline solar modules is a challenging task. In this work, we propose a weakly supervised learning strategy that only uses image-level annotations to obtain a method that is capable of segmenting cracks on EL images of solar cells. We use a modified ResNet-50 to derive a segmentation from network activation maps. We use defect classification as a surrogate task to train the network. To this end, we apply normalized  $L_p$  normalization to aggregate the activation maps into single scores for classification. In addition, we provide a study how different parameterizations of the normalized  $L_p$  layer affect the segmentation performance. This approach shows promising results for the given task. However, we think that the method has the potential to solve other weakly supervised segmentation problems as well.

**Index Terms**— crack detection, weakly supervised semantic segmentation, EL imaging, solar cell, normalized  $L_p$  norm

## 1. INTRODUCTION

The need of renewable energy sources is rapidly growing in times of climate change and world wide increasing energy consumption. Already today, a major part of renewable energy is obtained by photovoltaics (PV) [11]. The important role of solar power in the transition from conventional to renewable power sources enforces a thorough assessment of the quality of PV modules from production to operation. Therefore, an automatic detection and classification of certain defects is essential.

In this work, we propose a crack segmentation pipeline that is applied on different types of solar cells. Especially, the grainy and noisy structure of polycrystalline solar cells renders an automatic segmentation, classification or even localization of defects challenging (cf. Fig. 1). We use a publicly



**Fig. 1.** EL image of a polycrystalline solar cell with a crack in the right upper corner. Red box shows the area where the crack is located. Blue arrows point to the busbars. Green boxes show noisy areas, which are no crack regions.

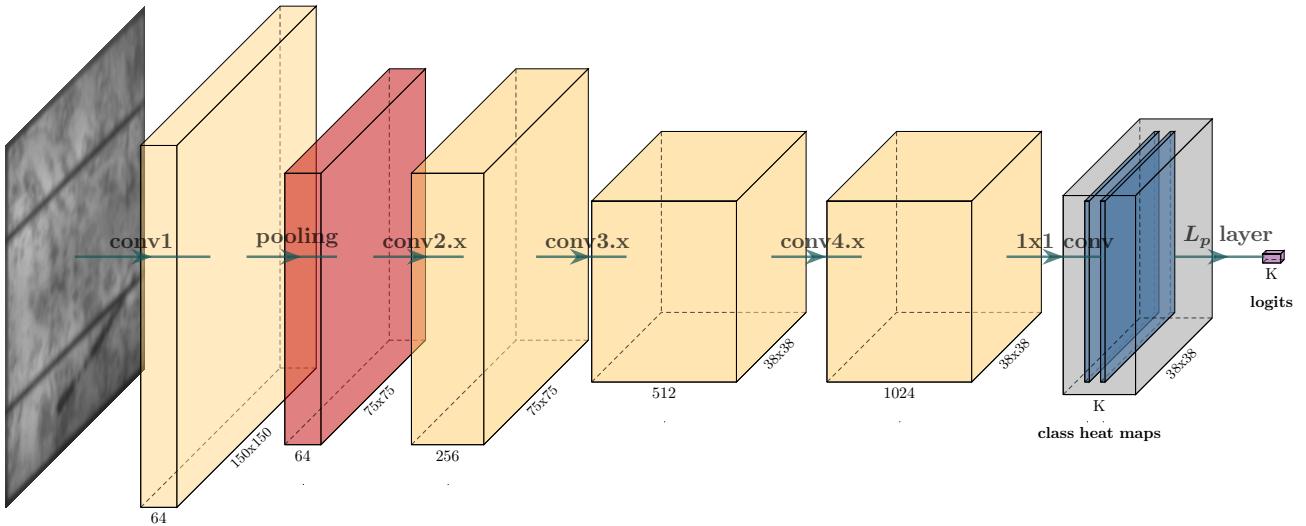
available dataset of electroluminescence (EL) images of solar cells. Since segmentation masks are unavailable and cost intensive to obtain, we resort to a weakly supervised segmentation approach using deep learning. This way, we learn a rough segmentation of the cell from image-level labels only. In addition, we systematically analyze the use of the normalized  $L_p$  norm with different values of  $p$  for the task of weakly supervised segmentation.

The main contributions of this work are:

- We developed a model that is capable of segmenting cracks on EL images of mono- or polycrystalline solar modules with one forward path.
- We analyzed the effect of different normalized  $L_p$  norms in case of weakly supervised semantic segmentation.

## 2. RELATED WORK

There are several algorithms which have been proposed to solve the problem of segmenting EL images of solar cells.



**Fig. 2.** Architecture of modified ResNet50. In contrast to the original ResNet50, this version is just downsampling in conv1 and conv3.x to retain more precise spatial information. Downsampling is performed via stride equal to 2 for the first convolutional layer of the scaled residual block.

In many cases, over-rejection of small cracks is a problem and, on the other hand, defects like dislocations are easily interpreted as cracks leading to a large number of false positives [1, 10, 14, 15, 16]. Since these works have shown that this problem is hard to solve with classical image processing methods, we apply a deep learning approach to solve it in a data driven way.

Weakly supervised segmentation with image-level labels has only rarely been addressed in research. As one of the pioneers, Oquab et. al. [12] recognized the potential of this learning strategy. They used a fully convolutional architecture with five convolutional blocks and added a global max-pooling layer at the end of the network. This approach results in a good localization of objects on the Pascal VOC 12 and MS COCO dataset. Later, Wei et. al. [17] extended the convolutional layers with different dilation rates to connect discriminative object regions with non-discriminative object regions. Huang et. al. [9] localize the seeds with a classification network and use these information cues for seeded region growing (SRG). In addition this SRG approach is extended with a segmentation network and a special seeding loss function. The results are promising. On the other hand, approaches based on Oquab et. al. [12] are attractive due to their simplicity and computational efficiency. In our architecture we use the normalized  $L_p$  norm described by Gulcehre et. al. [7] as a global pooling layer separately for each feature map. A similar pooling strategy has been used in the context of image retrieval by Radenovi et. al. [13].

### 3. MATERIALS & METHODS

#### 3.1. Training data and Preprocessing

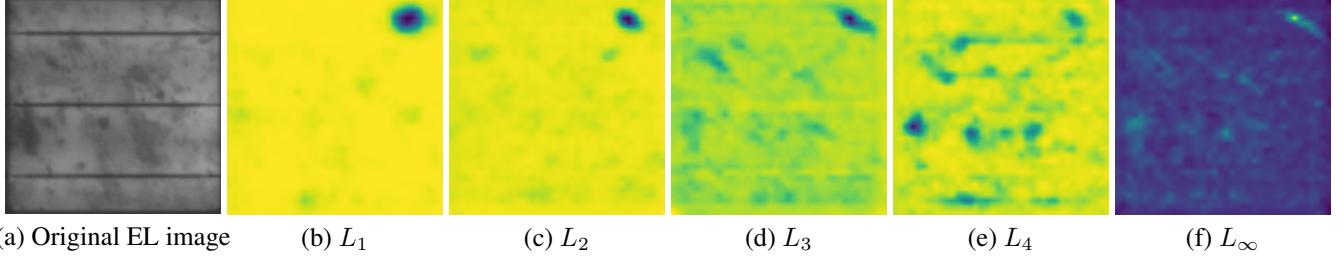
The dataset used for this work<sup>1</sup> consists of 2426 electroluminescence 8-bit grayscale images of solar cells with a resolution of  $300 \times 300$  pixels per image [3, 4, 5]. These images were extracted from 44 different solar modules. It includes expert annotations of the probability that a cell contains any kind of defect. We manually added image-level annotations of the defect type, since we are interested in cracks only. The dataset is class-imbalanced with much more non-crack samples than crack samples. It contains monocrystalline as well as polycrystalline cells.

In Fig. 1, an EL image of a polycrystalline solar cell is shown. The blue arrows point to the three vertical busbars which are visible on most of the samples and should not be segmented as crack regions. Since the configuration of busbars varies from cell type to cell type, the method needs to ignore the busbars reliably. Especially for polycrystalline cells, there are instances, where the structure of the silicon wafers is similar to the appearance of a crack (green boxes). Of course, neither of them should be segmented as cracks.

#### 3.2. Architecture

We start with a general classification architecture and modify it to obtain the segmentation result. The basis for this approach is a residual network with 50 layers [8]. Due to the

<sup>1</sup><https://github.com/zae-bayern/elpv-dataset>



**Fig. 3.** Original EL image and heat maps of crack layer for different normalized  $L_p$  norms. In the heat maps the dark areas are low values and the bright values are high values.

limited amount of data, best results are achieved with transfer learning. The weights are initialized with pretrained weights trained on ImageNet [6]. We use categorical cross entropy as loss function.

We apply four modifications to adapt the network to the given task:

1. To preserve spatial information throughout the forward path and increase the resolution of the segmentation masks, we remove the fully connected layer, the average pooling layer and one of the bottleneck blocks.
2. The last bottleneck block is changed to have a stride of 1, such that the resolution is not affected. To this end, the stride in the first convolutional layer is set to 1. To allow using the pretrained weights, we set the dilation rate of the following convolution to 2.
3. A  $1 \times 1$  convolutional layer with two feature maps is added to adapt the number of activation maps such that it matches the number of classes. In our case, the first feature map shows the activations for 'crack' and the other one shows the activations for 'non-crack'.
4. A normalized  $L_p$  norm layer is added to aggregate the activation maps into a single value for classification.

The input data from the feature maps is flattened to a finite set of input signals represented as vector  $\vec{x}$  with  $N$  elements. This vector is fed into the normalized  $L_p$  norm, which is defined as

$$L_p(\vec{x}) = \left( \frac{1}{N} \sum_{i=1}^N |y_i|^p \right)^{\frac{1}{p}}, \quad (1)$$

where  $y_i$  denote the  $N$  pixels of the corresponding activation map  $\mathbf{Y}$ .

Certain normalized  $L_p$  norms are equal to some types of pooling: For example,  $L_1$  is identical to average pooling and  $L_\infty$  is equal to max pooling [2, 7].

### 3.3. Segmentation task

The classification problem is used as a surrogate task to train the  $1 \times 1$  convolutional layer. At test time, this is a segmentation problem. Hence, we are only interested in the activations

Norm	$L_1$	$L_2$	$L_3$	$L_4$	$L_5$	$L_9$	$L_\infty$
<b>Precision</b>	0.90	0.82	0.87	0.83	0.54	0.58	0.77
<b>Recall</b>	0.77	0.78	0.80	0.67	0.70	0.63	0.77
<b>f1 score</b>	0.83	0.80	0.83	0.74	0.61	0.61	0.77
<b>Accuracy</b>	0.94	0.92	0.94	0.91	0.82	0.84	0.91

**Table 1.** Classification results for different  $L_p$  norms over the elpv-dataset. Precision, recall and f1 score measured on class 'crack'. Accuracy measured on all test data.

of conv5. These activations highlight the important regions during inference. The resolution of both heat maps is  $38 \times 38$  due to the downsampling of the ResNet-50. For the segmentation task, we only use the map that corresponds to the class 'crack'. In addition, we only apply the segmentation, when the classifications states that there is a crack on the input image. Then, the segmentation is simply done by pixel-wise comparing the heat map value with the half of the maximum value of the heat map

$$\text{segment}(y_i) = \begin{cases} 1, & \text{for } y_i > \frac{\max(\mathbf{Y})}{2} \\ 0, & \text{else} \end{cases} \quad (2)$$

If the compared value is above it is a crack region, otherwise it's background, shown in equation 2. The threshold is computed dynamically, since the activations are scaled arbitrarily.

Note that the actual segmentation is not the main purpose of this paper and a better segmentation can certainly be achieved by an advanced method.

## 4. RESULTS

First, we present classification results for different normalized  $L_p$  norms to show how the surrogate task performs using different values of  $p$ . Then, the heat maps and segmentation results are qualitatively evaluated to check whether good classification performances result in good segmentation results.

### 4.1. Classification

In Tab. 1, classification results for the class 'crack' are shown for different values of  $p$ . It turns out that  $L_1$  and  $L_3$  have

the best accuracy and f1 score during classification of the test samples.  $L_3$  has a slightly lower precision but higher recall than  $L_1$ . The norms  $L_2$ ,  $L_4$  and  $L_\infty$  have the same accuracy, but the f1 score differs strongly between them. The low f1 score of class 'crack' of  $L_4$  shows that the model favors the 'non-crack' class. The norms between  $L_5$  and  $L_9$  have low f1 scores and are therefore not taken into further consideration. Based on these classification results,  $L_1$  and  $L_3$  perform best.

## 4.2. Segmentation

Fig. 3 shows the results of one sample for different normalized  $L_p$  norms. These results differ from the classification results.  $L_1$  localizes the cracks very well, but it can't preserve the structure and leads to a blurry result.  $L_2$  has a similar behavior but with a finer clue for the location. The activation map of  $L_3$  gets the shape quite well. From  $L_4$  to  $L_9$  there is no region shown which can be interpreted as a crack. By contrast the activation map of  $L_\infty$  describes the crack very sharply. Hence it serves as a basis for the segmentation.

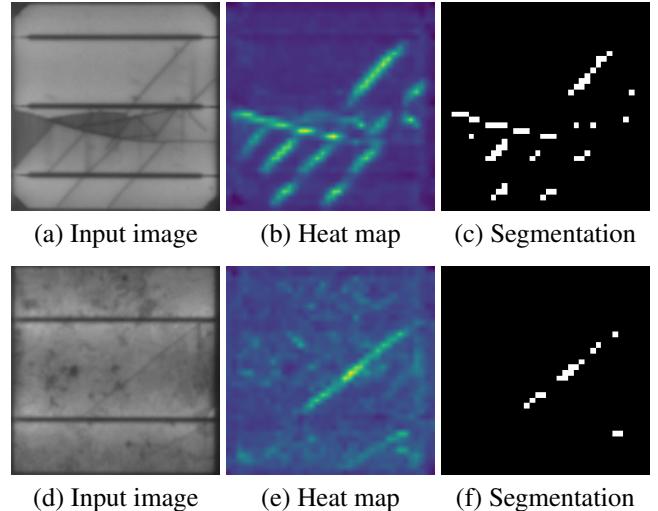
Interestingly, for  $L_1$  to  $L_4$ , the important regions contain low values in the activation maps and the mean of the activation map of class 'crack' is very high. For larger values of  $p$  this behavior is inverted, so that the important regions are represented as high values in the activation map. This is certainly explained by the fact that cracks cover only a small region of the entire image. To obtain a good classification score, a large number of pixels needs to be high in case of small  $p$ , which is achieved by the inversion of intensities.

In Fig. 4 the difference between heat maps for mono- and polycrystalline cells is highlighted. For EL images of monocrystalline cells, there is no noisy background and the structure of the crack is preserved nicely. In contrast, the heat maps for polycrystalline cells have a noisy background. Nevertheless, our proposed segmentation method gives good results in any case.

## 5. DISCUSSION

The results show that the choice of the best norm for a given problem is crucial for the quality of the weakly supervised segmentation results. While lower norms lead to blurred results, higher norms give very sharp results. From this behavior, we conclude that lower norms might be better suited for segmenting large, blob-like objects, whereas higher norms may be preferred for tiny objects with sharp edges.

For the given problem of segmenting cracks on solar cells, we see that  $L_\infty$  performs best, but  $L_3$  could be a good choice as well, if the foreground-background inversion is taken into account. In addition, we show that segmentation performance is not strictly related to classification performance.



**Fig. 4.** Results for input images of mono- and polycrystalline solar cells with normalized  $L_\infty$  norm

## 6. CONCLUSION

In this paper, we propose an effective technique to get a coarse segmentation of cracks from EL images of mono- and polycrystalline solar cells. We show that given only a small number of training samples and only image-level annotations, the method is able to reliably segment cracks on solar cells with just one forward pass through the network. We are confident, that this is an important contribution to an effective production and maintenance of solar power plants.

This application-oriented result is supplemented by a comprehensive study of the effect of normalized  $L_p$  normalization with different values of  $p$  on the result of weakly supervised segmentation methods. Thus, this method can serve as a basis for the design of new weakly supervised segmentation procedures.

## Acknowledgement

We gratefully acknowledge the Federal Ministry for Economic Affairs and Energy (BMWi: Grant No. 0324286D, iPV4.0) for funding.

## 7. REFERENCES

- [1] S. A. Anwar and M. Z. Abdullah. Micro-crack detection of multicrystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique. *EURASIP Journal on Image and Video Processing*, 2014(1):15, 2014.
- [2] Y.-L. Boureau, J. Ponce, and Y. Lecun. A theoretical analysis of feature pooling in visual recognition. *ICML*

- 2010 - Proceedings, 27th International Conference on Machine Learning*, pages 111–118, 11 2010.
- [3] C. Buerhop-Lutz, S. Deitsch, A. Maier, F. Gallwitz, and C. J. Brabec. A benchmark for visual identification of defective solar cells in electroluminescence imagery. In *35th European PV Solar Energy Conference and Exhibition*, 2018. accepted.
  - [4] S. Deitsch, C. Buerhop-Lutz, A. Maier, F. Gallwitz, and C. Riess. Segmentation of photovoltaic module cells in electroluminescence images. *Solar Energy*, 2018. submitted.
  - [5] S. Deitsch, V. Christlein, S. Berger, C. Buerhop-Lutz, A. Maier, F. Gallwitz, and C. Riess. Automatic classification of defective photovoltaic module cells in electroluminescence images. *Solar Energy*, 2018. submitted.
  - [6] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. ImageNet: A Large-Scale Hierarchical Image Database. In *CVPR09*, 2009.
  - [7] C. Gulcehre, K. Cho, R. Pascanu, and Y. Bengio. Learned-norm pooling for deep feedforward and recurrent neural networks. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 530–546. Springer, 2014.
  - [8] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
  - [9] Z. Huang, X. Wang, J. Wang, W. Liu, and J. Wang. Weakly-supervised semantic segmentation network with deep seeded region growing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7014–7023, 2018.
  - [10] W.-J. Lin, Y.-H. Lei, and C.-H. Huang. Automatic detection of internal defects in solar cells. In *Instrumentation and Measurement Technology Conference (I2MTC), 2011 IEEE*, pages 1–4. IEEE, 2011.
  - [11] U. Nations. *2018 Energy Statistics Pocket Book*. United Nations Publications, 2018.
  - [12] M. Oquab, L. Bottou, I. Laptev, and J. Sivic. Is object localization for free? - weakly-supervised learning with convolutional neural networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.
  - [13] F. Radenović, G. Tolias, and O. Chum. Cnn image retrieval learns from bow: Unsupervised fine-tuning with hard examples. In *European conference on computer vision*, pages 3–20. Springer, 2016.
  - [14] S. Spataru, P. Hacke, and D. Sera. Automatic detection and evaluation of solar cell micro-cracks in electroluminescence images using matched filters. In *Photovoltaic Specialists Conference (PVSC), 2016 IEEE 43rd*, pages 1602–1607. IEEE, 2016.
  - [15] D. Stromer, A. Vetter, H. C. Oezkan, C. Probst, and A. Maier. Enhanced crack segmentation (ecs): A reference algorithm for segmenting cracks in multicrystalline silicon solar cells. *IEEE Journal of Photovoltaics*, 9(3):752–758, 2019.
  - [16] D.-M. Tsai, S.-C. Wu, and W.-C. Li. Defect detection of solar cells in electroluminescence images using fourier image reconstruction. *Solar Energy Materials and Solar Cells*, 99:250–262, 2012.
  - [17] Y. Wei, H. Xiao, H. Shi, Z. Jie, J. Feng, and T. S. Huang. Revisiting dilated convolution: A simple approach for weakly-and semi-supervised semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 7268–7277, 2018.