Automated Detection of Solar Cell Defects with Deep Learning

Alexander Bartler, Lukas Mauch and Bin Yang
Institute of Signal Processing and System Theory
University of Stuttgart
Stuttgart, Germany
{alexander.bartler, lukas.mauch, bin.yang}@iss.uni-stuttgart.de

Michael Reuter
and Liviu Stoicescu
Solarzentrum Stuttgart GmbH
Stuttgart, Germany
{michael.reuter, liviu.stoicescu}@solarzentrum-stuttgart.com

***Abstract*—Nowadays, renewable energies play an important role to cover the increasing power demand in accordance with environment protection. Solar energy, produced by large solar farms, is a fast growing technology offering environmental friendly power supply. However, its efficiency suffers from solar cell defects occurring during the operation life or caused by environmental incidents. These defects can be made visible using electroluminescence (EL) imaging. A manual classification of these EL images is very time and cost demanding and prone to subjective inter-examiner variations. For a fully automated defect detection, we introduce a deep learning based classification pipeline operating on the EL images. This includes image preprocessing for distortion correction, segmentation and perspective correction as well as a deep convolutional neural network for solar defect classification with special emphasis on dealing with highly imbalanced dataset. The impact of minority oversampling and data augmentation on the system accuracy is investigated. The performance of our proposed classification pipeline is demonstrated by applying it to a real world dataset.**

***Index Terms*—solar cell classification, imbalanced data, deep learning, renewable energies, electroluminescence**

I. Introduction

Renewable energies are essential for future power supply. Besides wind and water energy, one of the most important technologies, supplying around two percent of the world's total power demand today, is solar energy [1]. It is mainly produced by huge solar parks with power outputs up to 1GW.

During their lifetime, the total power output decreases, mostly because of defects in the solar modules and their cells. An increasing number of environmental incidents such as storm and hail makes it even more important to analyze the solar panel condition and detect the occurred defects. An appropriate method is to use electroluminescence (EL) characterization under daylight conditions [2], [3]. Normal photography imaging is not able to reveal the solar cell defects. By means of EL imaging, it is possible to visualize defects like cracks and inactive cell areas to evaluate the cell quality and the overall module and solar park quality.

Manually analyzing these EL images is a very time consuming process because of the large amount of data. For example, in a 300MW park with 250W modules and 72 cells per module, 86.4 millions cell images need to be evaluated, which is infeasible in practice, see Table I. In addition, the quality of manual inspection of cell images depends on the experience of the examiner and is prone to subjective inter-examiner variations.

TABLE I: Number of cells in photovoltaic systems.

	rooftop		open space			
	private	industry	small	mid	large	record
MW	0.01	0.4	1	10	300	1,000
Modules	40	1,600	4,000	40,000	1,200,000	4,000,000
Cells	2,880	115,200	288,000	2,880,000	86,400,000	288,000,000

In order to enable a fast, low-cost and reliable evaluation of solar cells, we propose an automated defect detection, using a deep convolutional neural network (CNN) for the EL cell image classification. To estimate the power output of solar modules by using the sun's position, neural networks have already been applied with great success to detect power losses in solar modules [4]. Furthermore, the detection of solar modules in low-quality satellite photos by the application of convolutional neural networks with a high detection rate shows the potential of deep learning based methods to evaluate images of solar modules [5].

One difficulty in the training of the CNN in our case is the highly imbalanced dataset due to a typically small fraction of defect cells. Furthermore, this minority class shows large variations in the structure and position of the defects. They appear as fine structured small cracks or large dark areas in the cell image due to different physical causes. Further challenges are low resolution and low sharpness of the cell images caused by the EL imaging principle and the perspective image.

In this paper, we examine for the first time the feasibility of a CNN to detect cell defects by using EL images of solar cells. We present a signal processing pipeline for image preprocessing and classification, which will enable the automated evaluation of large solar parks in the future. Different methods to handle the dataset imbalance and the minority class variations are described. We investigate especially the impact of minority oversampling and data augmentation. We apply our pipeline to a real world dataset with promising results.

II. Methods

A. Pipeline

For the fully automated classification of solar cell defects, we developed a pipeline which processes the EL images of

ISBN 978-90-827970-1-5

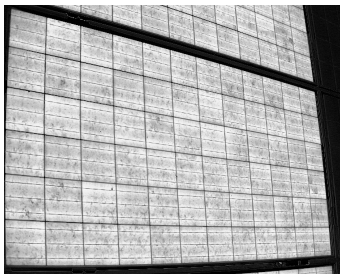

Fig. 1: EL image of a solar module.

solar modules. Fig. 1 shows one example of EL images of a solar module containing 6×12 cells. Since the EL images contain the intrinsic distortions caused by the camera lens, a first intrinsic calibration step is needed to correct the image distortions. This is done offline by using a calibration pattern and a standard calibration method [6].

Because the images are taken from a camera standing on the ground, a perspective transform as in Fig. 1 is present and has to be corrected as well. Therefore, the module in an EL image is segmented by means of a contour finding algorithm that extracts the raw contour of the module [7]. Since the perspective transform leads to an arbitrary quadrangle shape of the module, the RANSAC algorithm [8] is used to fit this geometrical shape into the raw contour ending up with the edge points of the module. The RANSAC algorithm helps to improve the accuracy of the estimated edge coordinates because of its robustness against outliers. With the knowledge of the width-to-height ratio of a module and the estimated edge coordinates, the perspective correction is performed. The cells are then extracted out of the corrected image by using the known number and arrangement of cells in a module (e.g. 6×12). Finally, the cells are classified by a CNN. An overview of the pipeline is shown in Fig. 2.

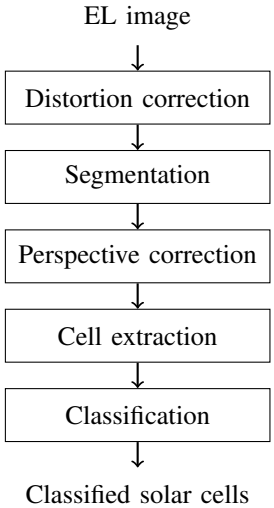

Fig. 2: Classification pipeline

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

Fig. 3: Cell images with (a) no defect, (b) micro crack defect, (c) large-scale defect, (d) defect and low resolution.

B. Images

The extracted gray cell images are resized to a fixed size of 120×120 pixels. Examples are shown in Fig. 3. The structure of the defects ranges from micro cracks to large dark areas due to different reasons. Different cell defects have different consequences, e.g. dark area leads to an immediate reduced power output while a crack can cause a reduced power output in the future. For this reason, many operators of solar parks wish an automated detection of defect cells and a further classification of defect cells into various defect categories in order to decide which solar modules (not cells) have to be replaced immediately or in the future. For a first feasibility study in this paper, we focus on the binary classification of good and defect cells. The multi-class classification problem requires a much higher number of labeled defect cells than currently available and thus will be addressed in the future.

Another result of the perspective imaging is the different resolution of cell images in the same module before resizing. In particular, the farthest cells may suffer from a too small original resolution (e.g. 50×50 pixels). It leads to blurred images where especially small defect structures may be hidden and hard to detect. This makes an automated analysis even more challenging.

C. CNN architecture

For the classification task, we adapted the VGG16 architecture [9] for a first feasibility study. More refined architectures like ResNet [10] will be studied in the future. We reduced the number of filters and the size of the fully connected layers to reduce the total number of parameters due to a smaller number of labeled training samples. Furthermore, we changed the output layer to fit our two-class problem. In Table II the architecture is briefly characterized.

We added a batch normalization between the convolution and the activation layer to speed up the training process and to be less sensitive to the initialization. The batch normalization also acts as regularizer which helps to avoid overfitting [11]. We used the exponential linear unit (ELU) function as activation [12]. All layers are regularized with the L2-norm to prevent overfitting and to enhance the generalization capability. Also dropout is used in the fully connected (FC) layers [13]. The weights are initialized with the He normal method [14].

ISBN 978-90-827970-1-5 © EURASIP 2018

TABLE II: Adapted VGG16 network

input (120×120 gray image)
conv3-8 conv3-8
maxpool
conv3-16 conv3-16
maxpool
conv3-32 conv3-3 2 conv3-32
maxpool
conv3-64 conv3-64 conv3-64
maxpool
conv3-64 conv3-64 conv3-64
maxpool
FC-128
FC-128
FC-2
softmax

convx-y:
convolution layer with x×x filter size and y feature maps

FC-x:
fully connected layer with x neurons

III. Oversampling and augmentation

A general problem in automated analysis of solar cell images is the data imbalance. Since EL images are taken from whole solar modules and since defect cells are relatively rare and mostly random distributed among modules, there is always a majority good class and a minority defect class of cells. A typical problem caused by the data imbalance is the bad classification accuracy of the minority class despite of a good performance on the majority class. In our case the false negatives (i.e. classify a defect cell as good) are much more critical than the false positives, because cell images which are classified as good will be never examined again by experts in fact of their large amount.

One way to combat the imbalance problem is to use a non-heuristic resampling method [15], [16]. An appropriate approach is to randomly oversample the minority class, i.e. using random copies of defect cell images to enlarge the minority class. This artificially balances the class distribution. Since oversampling duplicates images, it is prone to overfitting on the training data.

To reduce overfitting, data augmentation is another approach [17], [18]. It helps to improve the classification performance [19], especially for a small amount of training data. By data augmentation, especially when oversampling is applied before, a larger dataset is generated. One augmentation method to increase the diversity of the data is random horizontal and vertical image flipping.

The cell images extracted from the corrected module images suffer from small rotation, translation and shearing. To reach a better generalization for unseen data, the network should be robust to these transforms. We apply these transforms to the training images as a further data augmentation step. We identified the range of typical transform parameters in our dataset as in Table III.

All transforms, i.e. rotation, translation, shearing and vertical and horizontal flipping, are applied to each image after the oversampling step using random parameters within the given range in Table III.

TABLE III: Range of data augmentation parameters.

transformation	parameter
rotation	$\pm 10\,^{\circ}$
translation	$\pm 10\,\%$
shearing	$\pm 0.2\,\mathrm{rad}$

IV. Dataset and experiments

A. Dataset

For a first feasibility study, we use a dataset of 98,280 labeled cell images extracted from 1,366 module images. The labeling is done by solar cell and EL imaging experts. Some cells were excluded from the dataset because they have multiple defect types or they are shaded by surrounding objects like trees. The dataset was split into $90\,\%$ training and $10\,\%$ validation data. The splitting is done modulewise to generate a fully unknown validation set. The class distribution of the dataset is shown in Table IV. Only $3.4\,\%$ of the cells belong to the defect class. This indicates the high imbalance of our dataset.

TABLE IV: Class distribution of a dataset and partition into training and validation set.

class	training	validation
good	85513	9474
defect	2915	378

B. Experiments

For the implementation of the image preprocessing steps, we used OpenCV [20]. To implement and evaluate the CNN architecture, the minority oversampling and the data augmentation, we used the Keras framework [21].

To normalize the input data, the mean gray value of the training set was subtracted from each sample of the training and validation set. We used a dropout ratio of 0.7 in the FC layers and a L2-regularization constant of 0.01 in all layers. The cost function is the cross-entropy loss.

The adaptive sub-gradient method Adagrad [22] was chosen as optimizer. It takes the updates of earlier iterations into account and helps the CNN to train on the minority class with its wide distribution. The initial learning rate is set to 0.01 and was divided by 10 after 20 epochs to improve the convergence. The network was trained with a batch size of 512 over 40 epochs.

To investigate the impact of minority oversampling and data augmentation to our dataset, we compare the following three experiments.

(a) Oversampling and no data augmentation

ISBN 978-90-827970-1-5 © EURASIP 2018

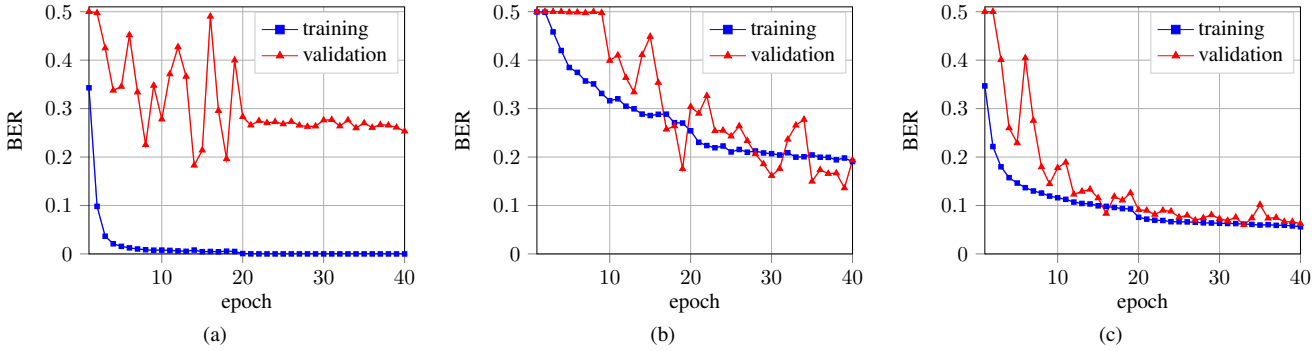

Fig. 4: Results of training and validation BER (a) oversampling, no augmentation, (b) no oversampling, augmentation, (c) oversampling, augmentation.

(b) No oversampling and data augmentation
(c) Oversampling and data augmentation

When the minority oversampling is applied, the minority class is enlarged by using multiple copies of defect cell images. The number of good and defect cell training images have the ratio 2:1 after oversampling. If the data augmentation is used, the defect and good cell images are flipped and transformed randomly according to Table III to enlarge the diversity of the training images. The training set remains imbalanced. The transforms are applied to the input images randomly right before each epoch. In the last experiment (c), both oversampling and data augmentation are applied leading to a more balanced training dataset (2:1) with enlarged diversity.

To compare the results, we focus on the balanced error rate (BER), the false positive rate (FPR) and the false negative rate (FNR), i.e classifying good/defect cells as defect/good, respectively. The main target is to keep the FNR as low as possible while having a low BER, because as defect classified cells can be checked by experts and as good classified cells will never be examined again.

V. Results

The impact of the class imbalance is clearly shown in Fig. 4 and Table V. The trained network with minority oversampling or data augmentation alone fails to correctly detect defects as shown by the FNR of 50.26% or 38.89% while most of the good cells are classified correctly. This is caused by the imbalanced dataset where the CNN suffers from overfitting and tends to decide for the majority of good cells.

By using only oversampling, the network performs quite well on the training set as indicated by the fast converged and low training BER (see Fig. 4a). However, the result on the validation set is poor. This means, the trained network is overfitted. The oversampling helps to train on both classes, but due to low variations with multiple copies of the minority class in the training set, the network is not able to generalize well.

Using data augmentation without oversampling helps to reduce the FPR and the FNR, but still a large value of FNR causes the high BER of 19.57% (see Fig. 4b). The data augmentation, the small random variations of the cell images, enhances the diversity of the training images, but not the number of minority training samples. Hence the dataset remains imbalanced.

When we combine both methods, the overfitting to the training set caused by the oversampling is suppressed by the data augmentation and the oversampling helps the network to train on the minority class. The training and validation BER decrease relatively fast and converge to almost the same value of around 7% with a low variance (see Fig. 4c). This is caused by the better classification capability for the class of defect cells, indicated by an FNR improvement from 50.26% and 38.89% to 12.96%. Since minimizing the FNR is more important than avoiding misclassification of good cells, we can tolerate the slightly increased FPR.

TABLE V: Comparison of the results on the validation set and the ratio between good and defect training cell images.

method	BER	FNR	FPR	ratio
oversampling	25.40%	50.26%	0.53%	$2:1$
augmentation	19.57%	38.89%	**0.25%**	$9:1$
both	**$7.73\ \%$**	**$12.96\ \%$**	2.50%	$2:1$

VI. Conclusion

In this paper we demonstrated a first feasibility study for the detection of defect solar cells. We succeeded to extract cells from perspective solar module images and classify the cell images by using a CNN in a fully automated way. In particular, we studied the data imbalance problem and compared different methods to address the issue. In a real experiment, we achieved a BER of 7.73% for the binary classification problem.

In the future, we will address the more challenging multi-class classification problem (different defect classes) with an even more serious imbalanced dataset. More advanced data augmentation techniques (e.g. generative adversarial network (GAN) [23]), refined CNN architectures (e.g. ResNet [10]) and other learning strategies (e.g. ensemble learning) will be studied to address the issue.

ISBN 978-90-827970-1-5 © EURASIP 2018

REFERENCES

[1] A. Jäger-Waldau, "Snapshot of Photovoltaics - March 2017," *Sustainability*, vol. 9, 05 2017.

[2] L. Stoicescu, M. Reuter, and J. H. Werner, "Daylight Luminescence for Photovoltaic System Testing," in *Proc. 22nd International Photovoltaic Science and Engineering Conference*, Hanghzhou, China, 2012.

[3] T. Kropp, M. Berner, L. Stoicescu, and J. H. Werner, "Self-Sourced Daylight Electroluminescence From Photovoltaic Modules," *IEEE Journal of Photovoltaics*, vol. 7, no. 5, pp. 1184–1189, Sept 2017.

[4] K. Jazayeri, S. Uysal, and M. Jazayeri, "Determination of power losses in solar panels using artificial neural network," in *2013 Africon*, Sept 2013, pp. 1–6.

[5] V. Golovko, S. Bezobrazov, A. Kroshchanka, A. Sachenko, M. Komar, and A. Karachka, "Convolutional neural network based solar photovoltaic panel detection in satellite photos," in *2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 1, Sept 2017, pp. 14–19.

[6] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov 2000.

[7] S. Suzuki and K. Abe, "Topological structural analysis of digitized binary images by border following." *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 1, pp. 32–46, 1985. [Online]. Available: http://dblp.uni-trier.de/db/journals/cvgip/cvgip30.html#SuzukiA85

[8] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981. [Online]. Available: http://doi.acm.org/10.1145/358669.358692

[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014. [Online]. Available: http://arxiv.org/abs/1409.1556

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[11] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *CoRR*, vol. abs/1502.03167, 2015. [Online]. Available: http://arxiv.org/abs/1502.03167

[12] D. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," *CoRR*, vol. abs/1511.07289, 2015. [Online]. Available: http://arxiv.org/abs/1511.07289

[13] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014. [Online]. Available: http://jmlr.org/papers/v15/srivastava14a.html

[14] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," *CoRR*, vol. abs/1502.01852, 2015. [Online]. Available: http://arxiv.org/abs/1502.01852

[15] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *SIGKDD Explor. Newsl.*, vol. 6, no. 1, pp. 20–29, Jun. 2004. [Online]. Available: http://doi.acm.org/10.1145/1007730.1007735

[16] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, Sept 2009.

[17] P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings.*, Aug 2003, pp. 958–963.

[18] D. C. Ciresan, U. Meier, L. M. Gambardella, and J. Schmidhuber, "Deep big simple neural nets excel on handwritten digit recognition," *CoRR*, vol. abs/1003.0358, 2010. [Online]. Available: http://arxiv.org/abs/1003.0358

[19] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: when to warp?" *CoRR*, vol. abs/1609.08764, 2016. [Online]. Available: http://arxiv.org/abs/1609.08764

[20] Itseez, "Open source computer vision library," https://github.com/itseez/opencv, 2015.

[21] F. Chollet *et al.*, "Keras," https://github.com/fchollet/keras, 2015.

[22] J. Duchi, E. Hazan, and Y. Singer, "Adaptive subgradient methods for online learning and stochastic optimization," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2010-24, Mar 2010. [Online]. Available: http://www2.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-24.html

[23] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," *ArXiv e-prints*, Jun. 2014.

ISBN 978-90-827970-1-5 © EURASIP 2018