Solar Energy 201 (2020) 453–460

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener

Deep learning based automatic defect identification of photovoltaic module using electroluminescence images

Wuqin Tang[a], Qiang Yang[a,b,*], Kuixiang Xiong[a], Wenjun Yan[a]

[a] College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China
[b] Zhejiang Lab, Hangzhou 310000, China

ARTICLE INFO

Keywords:
Electroluminescence Images
Convolution neural network
Automatic defect classification
Generative adversarial network

ABSTRACT

The maintenance of large-scale photovoltaic (PV) power plants is considered as an outstanding challenge for years. This paper presented a deep learning-based defect detection of PV modules using electroluminescence images through addressing two technical challenges: (1) providing a large number of high-quality Electroluminescence (EL) image generation method for the limit of EL image samples; and (2) an efficient model for automatic defect classification with the generated EL image. The EL image generation approach combines traditional image processing technology and GAN characteristics. It can produce a large number of EL image samples with high resolution using a limited number of samples. Then, a convolution neural network (CNN) based model for the automatic classification of defects in an EL image is presented. CNN is used to extract the deep feature of the EL image. It can greatly increase the accuracy and efficiency of PV modules inspection and health management in comparison with the other solutions. The proposed solution is assessed through extensive experiments by using the existing machine learning models, VGG16, ResNet50, Inception V3 and MobileNet, as the comparison benchmarks. The numerical results confirm that the proposed deep learning-based solution can carry out efficient and accurate defect detection automatically using the electroluminescence images.

1. Introduction

In the past decade, solar photovoltaic (PV) energy as clean energy has received tremendous attention and experienced a dramatically rapid development across the world. The rapid increase of PV deployment, including both centralized PV farms and distributed PV generation (e.g., roof-top panels), is mainly driven by the PV technological advances and the decreased cost. The recent statistics (Frankfurt School-UNEP Centre/BNEF, 2019) shows that the levelised cost of electricity (LCOE) of PV generation has dropped about 81% in the past 10 years. The International Energy Agency (IEA) reports that the world's total renewable-based power capacity will grow by 50% between 2019 and 2024 and the solar PV accounts for 60% of the capacity addition (IEA, 2019). The total global PV capacity is estimated to increase from about 593.9 GW (GW) in 2019 to about 1500 GW in 2030 following significant capacity additions by China, India, Germany, the US and Japan. In particular, the solar-based energy consumption is expected to reach 33.0 Mtoe by 2023 in China (International Energy Agency, 2018). The crystalline silicon (c-Si), both mono and poly, is the main type of it, which accounts for over 97% of the market share (Lupangu and Bansal, 2017). Other technologies such as the Copper Indium Gallium diselenide (CuInSe2), Cadmium Telluride (CdTe) and amorphous-Si (a-Si) are applied in a small proportion (Birkmire, 2001) and c-Si is the dominant material for the PV industry (Parretta et al., 2005; Burrows and Fthenakis, 2015; Pazheri et al., 2014). The statistics (Burrows and Fthenakis, 2015) reported that about 362 GWp of c-Si modules has been installed worldwide and the c-Si modules have received more attention in recent years due to better performance on reliability assessment. However, even though the PV installed capacity is huge, there still lacks efficient and intelligent solutions to address the maintenance of large-scale PV based power generation infrastructures.

In reality, the PV modules can be damaged during the transportation, installation and operation processes. Potential-induced degradation (PID) can directly cause significant power losses (greater than30%) to silicon PV modules, especially in a humid and hot environment (Chen et al., 2017; Kang et al., 2015; Luo et al., 2016). The current-voltage (IV) curve is widely applied in the realm of PV module testing. However, some defects, e.g., degrading fingers or deteriorating contact dots in metal-wrap-through (MWT) solar cells, just cause negligible variations in IV data, that can be hardly detected until it reaches severe stages of degradation (Chegaar et al., 2006; Carson, 2008; Sayigh, 2015; Bouzidi et al., 2007; Rosch et al., 2012). The Infrared (IR)

* Corresponding author at: College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China.
E-mail address: qyang@zju.edu.cn (Q. Yang).

https://doi.org/10.1016/j.solener.2020.03.049
Received 6 November 2019; Received in revised form 13 March 2020; Accepted 14 March 2020
Available online 17 March 2020

imaging is a contactless technique for PV modules testing. The defected modules can generate the hot spot effect, which can be imaged by IR camera due to the relatively high temperature in the defected area. However, such a method can only detect the anomalous temperature and cannot effectively distinguish the defect types. Also, some defects, e.g., micro-crack, can be hardly detected as no abnormal temperature can be observed. Also, the resolution of IR images is generally low, which may directly degrade the performance of defect detection and analysis.

On the other hand, the Electroluminescence (EL) imaging is another non-destructive and contactless technique mainly for quality control of PV modules through defect detection and analysis. In general, the forward bias current needs to be injected into the modules that cause EL emission at a wavelength of 1150 (Coello, 2011; Yongqing et al., 2013; Fuyuki et al., 2007) to obtain the EL images. A silicon Charge-coupled Device (CCD) sensor unit is adopted to capture the images in front of the PV modules (Köntges et al., 2014). The obtained EL images with high resolution enable the detection of PV module defects, e.g., micro-cracks (Breitenstein et al., 2011). Even though EL inspection needs some time and experienced specialists, it has become the main method for defect detection of PV cells due to its excellent performance. In this paper, an automatic method is proposed for solving the limits.

The defects in solar modules can be mainly classified into three categories (Breitenstein et al., 2011): (a) micro-crack; (b) break; (c) finger-interruption. The finger-interruption does not necessarily cause a reduction in energy conversion. The micro-crack defects can degrade the PV generation and even lead to module failure. Fig. 1 illustrates these defects in both polycrystalline and monocrystalline PV cells (Deitsch et al., 2018).

In recent years, deep learning has been widely considered as one of the most powerful tools for pattern recognition and classification, and has been adopted in many application domains, e.g., autopilot, robotic automation, and health care. In the literature, many forms of deep learning-based methods have been developed for automatic classification, e.g., AlexNet (Krizhevsky et al., 2012), and VGG (Simonyan and Zisserman, 2015), GoogleNet (Szegedy et al., 2015), and ResNet (He et al., 2016). However, the classification performance of these approaches firmly relays on the training process based on a sufficient number of image samples. This imposed a significant challenge to adopt the deep learning based defect detection and analysis for PV modules. In most recent years, the generative adversarial network (GAN) (Goodfellow et al., 2014) has been developed to provide a viable solution to tickle such a problem through the generation of new images through extracting the features of the input image dataset. There is a collection of GAN-based models, e.g., DCGAN (Alec et al., 2015), WGAN (Arjovsky and Bottou, 2017; Martin et al., 2017), WGAN-GP (Gulrajani et al., 2017), as well as other variants are available. The previous investigations and solutions confirm the feasibility of the deep learning-based automatic detection and analysis of PV modules.

In recent years, the PV module defect detection and analysis using EL images have received much attention. The authors in (Tsai et al., 2013) proposed a method based on independent component analysis (ICA) for defect detection using EL images, and the accuracy of 93.4% was obtained. The authors in (Spataru et al., 2016) adopted the micro-crack detection and evaluation method to a 60-cell mc-Si module with extensive cell cracks, which can detect over 90% of cracks larger than 3 cm and the micro-crack location map is also detected to visualize if and where the type of micro-cracks is. To the best of our knowledge, micro-crack is just one type of defect. It is not enough for condition evaluation and the length of some micro-cracks is shorter than 3 cm. In (Sovetkin and Steland, 2018), a Hough transform-based approach was developed that can well identify and extract the cells in PV modules through correct the image rotation and distortions with complex backgrounds. However, the detection of PV defects was not investigated. In (Deitsch et al., 2018), the authors proposed two automatic machine-learning enabled defect detection solutions based on convolution neural network (CNN) and Support Vector Machine (SVM), respectively. These two approaches were trained using 1968 PV cell images and can analyze the PV cells (with a defect or defect-free) with an accuracy of 88.42% and 82.44%, respectively. However, these solutions can merely identify the PV cells with defects without the determination of defect types. The study in Luo et al. (2016) proposed a solution for finger interruption detection in PV cells based on the line scan photoluminescence (PL) imaging. This method provided a direction for finger detection. However, presently PL imaging technology has not been widely adopted in the PV energy sector. In (Mayr et al., 2019), a defect detection method using ResNet50 was developed for segmenting cracks on EL images with only a small number of training samples. However, this work only considered a limited number of defect types and more defects need to be investigated. In (Demirci et al., 2019), a transfer learning-based model for defect detection was developed. The transfer learning technique is considered an efficient method in the case of limited availability of training samples. In general, the transfer learning models can be used to process the EL images with a simple background, but the accuracy can be significantly degraded in the case of sophisticated backgrounds. In (Qian et al., 2018), a micro-cracks detection solution based on conventional digital image processing was presented. However, the learning-based models are considered more robust and efficient for pattern recognition of images than such conventional methods. In (Akram et al., 2019; Dhimish and Holmes, 2019), the CNN-based models were proposed for classification based on the data augmentation. However, such methods can suffer from the over-fitting phenomenon, especially under the conditions of training samples created by the geometric deformation. A structure-aware-based method for defect detection is proposed in Chen et al. (2020), which analyzes the structure feature of defects and noise. Then, a model is designed to filter the noise in EL image and in Chen et al. (2019), a novel steerable evidence filter is proposed to create the crack saliency map and a segmentation method is employed to extract the complete crack. However, the robustness of the models in Chen et al. (2020, 2019) is a limitation and it will be better if the method could combine with the learning-based model.

In summary, it can be observed that the existing solutions have limitations and the efficient defect detection and analysis of PV modules with EL images still demand further research exploitation from many aspects: (1) an efficient method for data augmentation; (2) more forms of module defects need to be considered; and (3) the detection accuracy and robustness of the machine learning models need to be

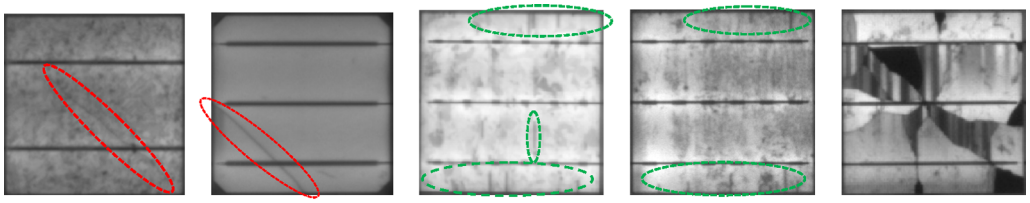

Fig. 1. Different forms of defects in PV cells: (a) micro-crack in polycrystalline silicon; (b) micro-crack in monocrystalline silicon; (c) finger-interruption in monocrystalline silcon; (d) finger interruption in polycrystalline silicon; (e) break.

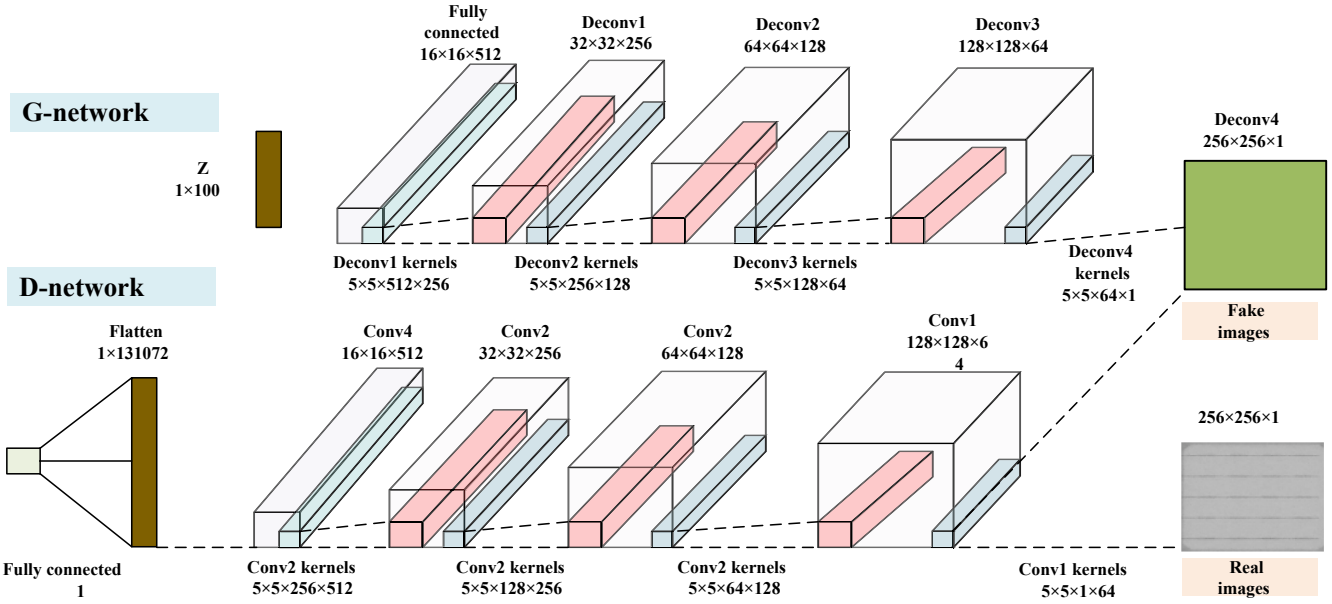

Fig. 2. Structure of GAN model.

further improved. To this end, this paper investigates a deep learning-based automatic defect detection solution for photovoltaic modules using the electroluminescence images. The main technical contributions made in this paper are summarized as two-fold:

(1) An efficient joint data augmentation approach combing the image alternation and GAN model is developed for EL images. It is confirmed that the additional augmented images can well tickle the issue of insufficient availability of data samples and significantly improve the performance of the machine learning models.
(2) A CNN based model for PV module defect detection and classification is developed based on the availability of EL images. The proposed solution is extensively assessed and validated through experiments by using the existing machine learning-based solutions, VGG16, ResNet50, Inception V3 and MobileNet, as the comparative benchmarks. Also, the performance with different model parameters is examined through a comparative study.

The remainder of the paper is organized as follows: Section 2 presents the training dataset augmentation approach; Section 3 provides the details of the developed CNN based automatic defect classification model; In Section 4, the performance of the proposed solution is evaluated through extensive experiments and the numerical results are presented and discussed; finally, the conclusive remarks and future work are given in Section 5.

2. The training dataset augmentation approach

In this work, an EL image dataset augmentation approach is developed to provide sufficient images for the machine learning training process. The proposed approach combines the conventional image processing algorithm and a GAN approach.

2.1. Conventional image processing method for image alternation

Based on the available EL images, the conventional image processing approaches can be used to produce additional images through translation and scaling. In this work, the number of generated image samples based on such a conventional processing approach is set 4 times the number of initial image samples. The training samples are scaled no more than 1%. The absolute value of the rotation range is less than 3°. The training samples are set to rotate by 90°, 180°, and 270° as well as vertical and horizontal mirror transformation in a random fashion.

The data augmentation based on the traditional image processing method is convenient with fast computation speed and the demand for computing hardware is low. However, the major deficiency is that those generated images have a minor difference between each other if many images are generated and it can lead to overfitting.

2.2. GAN-based model for data augmentation

The adopted GAN model consists of two parts: the generator network (G-network) and the discriminator network (D-network), as illustrated in Fig. 2.

There are four transposed convolution layers and a fully connected layer in the G-network. The fully connected layer takes a noise distribution as input and changes the 100-dimensional vector into a 16 * 16 * 512 tensor and generates the fake images (256 * 256 * 1) by extracting the local features with 256, 128, 64, and 1 transposed

Table 1
Parameters of GAN model.

G-network	Output shape	Parameters	D-network	Output shape	Parameters
Full connected	16 × 16 × 512	13,238,272	Convolution layer 1#	128 × 128 × 64	1,664
Deconvolution layer 1#	32 × 32 × 256	3,277,056	Convolution layer 2#	64 × 64 × 128	204,928
Deconvolution layer 2#	64 × 64 × 128	819,328	Convolution layer 3#	32 × 32 × 256	819,456
Deconvolution layer 3#	128 × 128 × 64	204,864	Convolution layer 4#	16 × 16 × 512	3,277,312
Deconvolution layer 4#	256 × 256 × 1	1,601	Flatten	131,072	0
			Full connected	1	131,073

Table 2
Initialized hyper-parameters of the GAN.

Hyper-parameters	Value
Batch size	36
Epoch	1000
Learning rate	0.0002
Decay	0.5

convolution kernels.

The D-network has four convolution layers and a fully connected layer for discriminating the real or fake images. The parameters of the D-network and G-network structure are given in Table 1.

In this paper, the proposed model establishes a G-network, which takes noise as input and output a fake image. The generated images are distinguished in D-network. The difference between the generated images and real images are measured by the distance of the Earth-Mover (EM) and the cost function of G-network and D-network are described in (1) and (2), respectively:

$$G_{loss} = -E_{x\sim P_g}[f_w(x)] \tag{1}$$

$$D_{loss} = E_{x\sim P_g}[f_w(x)] - E_{x\sim P_r}[f_w(x)] \tag{2}$$

where f_w denotes the D-network model and the input x comes from the image generated or the real images. Root Mean Square prop (RMSprop) is chosen as the optimization algorithm to train the model. To make $-D_{loss}$ closer to Wasserstein distance, the D_{loss} should be minimized. The initial hyper-parameters of the GAN is shown in Table 2.

Compared with the conventional image processing method, GAN based needs more computing time and has relatively higher computing hardware demand. The generated EL images with different defects using the GAN model are illustrated in Fig. 3. The Fréchet Inception Distance (FID) and the maximum mean discrepancy (MMD) are adopted to evaluate the performance of WGAN, as shown in (3) and (4). Here, in total 500 images are randomly selected from the real EL dataset, and the datasets are generated based on WGAN and DCGAN, respectively. From the Table 3, it is obvious that WGAN outperforms DCGAN for data augmentation of EL images.

$$FID = ||mu\_1 - mu\_2||^2 + Tr(C\_1 + C\_2 - 2 * sqrt(C\_1 * C\_2)) \tag{3}$$

$$MMD^{[2]}(P_r, P_g) = E_{x_r P_r, x_g P_g}[|| \sum_{i=1}^{n1} k(x_r) - \sum_{i=2}^{n2} k(x_g)||] \tag{4}$$

Here, “mu_1” and “mu_2” refer to the feature-wise mean of the real and generated images. C_1 and C_2 represent the covariance matrix for the real and generated feature vectors, often referred to as sigma. P_g and P_r are the distribution of generated samples and real samples, respectively.

2.3. Combined data augmentation method

Two effective data augmentation methods are presented in Sections 2.1 and 2.2, respectively, that both have own advantages and disadvantages. The conventional method can generate a new image with less computation time and lower hardware demand in comparison with GAN. On the contrary, GAN can obtain those images with a greater difference than the conventional method that may easily lead to overfitting.

Table 3
Performance assessment of WGAN and DCGAN model.

Dataset	MMD	FID
DCGAN	0.835	106.32
WGAN	0.793	86.834
Real dataset	0.649	65.389

The convolutional layers in CNN cannot obtain the rotation invariance feature for an image. The images with different directions or geometric deformation can be treated as new samples. However, the benefit of such an image augmentation solution is constrained due to the limited image variability.

In contrast, GAN based solution can extract the deep image features to enable image augmentation. These augmented images are significantly different that can improve the performance of the CNN model through the training process. It should be noted that an appropriate number of images should be adopted during the CNN model training process to prevent overfitting.

In addition, GAN-based image generation generally exhibits randomness, and hence the generated images may not contain the rotation, translation and scaling features. To make the data augmentation more efficient under the condition of limited image availability, the joint approach of combining GAN model and conventional image processing algorithm for image augmentation is used, i.e., GAN model is firstly used to generate the images and then the conventional image processing is used to enlarge the size of the sample image dataset.

3. CNN based automatic defect classification using EL images

This section presents the details of the proposed CNN based automatic PV module defect detection using EL images. In CNN based model, the operation of alternate convolution and sub-sampling are performed and the output is flatted as the input of a fully connected layer, as adopted in Goodfellow et al. (2014) and Alwani et al. (2016). To make the features acquired previously correspond to the defect type, the softmax (Liu et al., 2016) is used as the last layer in this work.

3.1. Defect classification model

In this work, the adopted CNN structure is illustrated in Fig. 4, consisting of eight learned layers: five convolution layers (c1–c5) and three fully connected layers (f1, f2, and output). The five convolution layers are used to extract features from EL images by using different kinds of kernels and then the features are sent to a fully connected layer. Finally, the output of a fully connected layer is made available to a 4-way classifier, which adopted the non-linear Rectified Linear Units (ReLU) (Jiang et al., 2018) and the features can be classified into 4 different defect labels. To further reduce computational complexity and hardware requirements, the second and fifth convolution layer are followed by a max-pooling layer. The detailed parameters of the adopted CNN in this work are shown in Table 4.

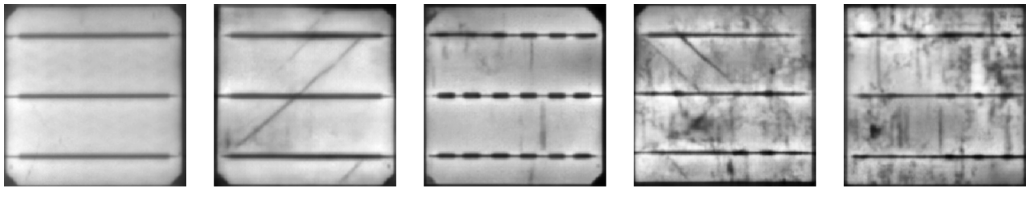

Fig. 3. EL images created by the proposed GAN-based augmentation approach (a) defect-free cell; (b) micro-crack in monocrystalline silicon; (c) finger-interruption in monocrystalline silicon; (d) micro-crack in polycrystalline silicon; and (e) finger-interruption in polycrystalline silicon.

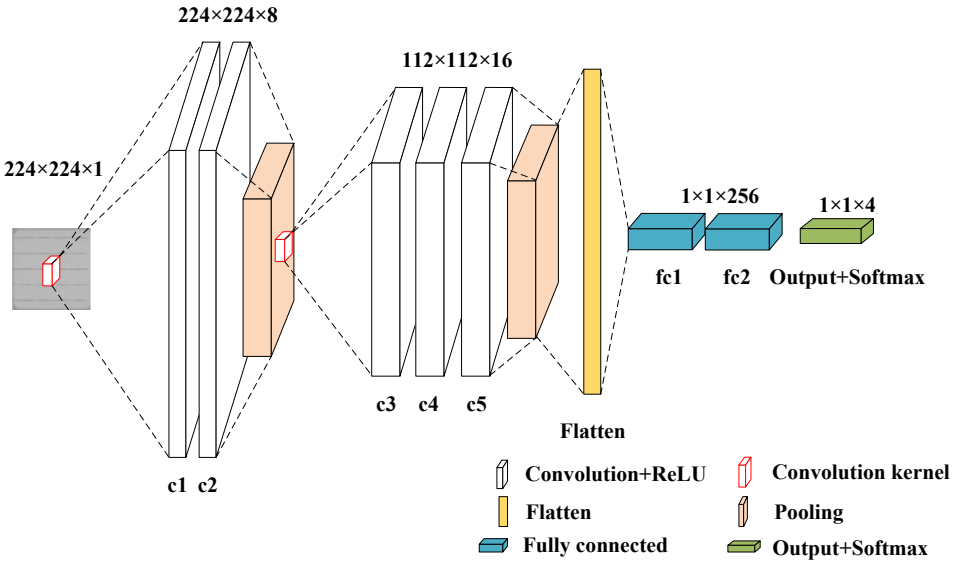

Fig. 4. Architecture of CNN based model for defect classification.

Table 4
Parameters of CNN based model for defect classification.

CNN	Output shape	Parameters
Input	224 × 224 × 1	0
Convolution layer 1	224 × 224 × 8	208
Convolution layer 2	224 × 224 × 8	1608
Max pooling	112 × 112 × 8	0
Convolution layer 3	112 × 112 × 16	3216
Convolution layer 4	112 × 112 × 16	6416
Convolution layer 5	112 × 112 × 16	6416
Max pooling	56 × 56 × 16	0
Flatten	50,176	0
Fully connected layer 1	256	12,845,312
Fully connected layer 1	256	65,792
Output	4	1028

3.2. Training process

The dataset of EL images is expressed as $\{(x_1, y_1), (x_2, y_2) \ldots, (x_n, y_n)\}$, where x is the EL image and y is the corresponding defect label, n is the number of EL images. To describe the degree of convergence of functions, a cross-entropy loss function is adopted in this work as expressed in (5).

$$E^{\mathrm{n}} = -\sum_{\mathrm{k}} t^{n}{}_{k} \log(y^{n}{}_{k}) \tag{5}$$

where y_k^n is the label of the image, the t_k^n is the predicted value of the CNN model, n is the number of EL images and m is the number of defect class. In this paper, the network is trained with the dataset generated from the combined method proposed in part 2.3.

During the training process, the weights in the network are constantly updated based on the stochastic gradient descent with momentum. The model training parameters are given in Table 5.

Table 5
Training parameters of the developed defect detection model.

Description	Symbol	Value
Batch size	b_s	64
Learning rate	η	1e-5
Momentum	ν	0.9
Weight decay	λ	0.01
Iteration times	T	500,000

4. Experimental assessment and numerical results

This section evaluates the performance of the proposed defect detection solution through a comparative study with existing machine learning models. In this work, the obtained dataset consists of the EL images representing four different types of module conditions: defect-free, micro-crack, break, and finger-interruption. The numerical results are presented to verify the effectiveness of the proposed solution and also the performance with different model parameters is extensively analyzed.

4.1. Experimental settings

In this work, the initial EL image dataset is obtained from both the public domain (Buerhop-Lutz et al., 2018) and the private dataset provided by JinkoPower Company (Shanghai, China).[1] The dataset consists of in total 1800 EL images, and 450 images for each type of defect condition (i.e. micro-crack, finger-interruption, break or defects-free). For each defect condition, 150 EL image samples are selected as the test dataset and 300 images are used to generate 2000 additional high-quality EL images by the use of the proposed GAN model. The produced EL images based on the GAN model are further augmented four times using the conventional image processing method, and 80% and 20% of the augmented images are used for the training dataset and validation dataset, respectively. As a result, the number of EL images of the training, validation and test dataset is shown in Table 6.

In these experiments, the parameters of GAN and CNN models are given in Table 1, Table 2, Table 4 and Table 5, respectively. The GAN and CNN based model is implemented in python based on TensorFlow. All these models are executed using a computer with 4 NVIDIA TITAN GPUs.

4.2. Training process assessment

The evaluation of a method for EL image sample augmentation can be judged directly by the quality of the generated EL image and further tested by the accuracy of the defects classification. Here, the performance of classification is evaluated. The accuracy can be output by the proposed CNN-based model and the convergence speed can be shown with the training time to the steady-state.

[1] The private image dataset is available upon request after permission.

Table 6
The number of EL images for each dataset.

Dataset (consisting of four types of defect conditions)	Dataset size (number of EL images)
Training set	1600 × 4 × 4
Validation set	400 × 4 × 4
Test set	150 × 4

Fig. 5 presents the accuracy and loss performance of the training dataset and validation set in details against the number of epochs during the training process. It can be seen that the accuracy increases rapidly and the loss decreases significantly for both datasets along with the increase of the number of epochs, and finally the training process converges within around 1200 epochs. The mean accuracy of about 100% and 83% can be obtained for the training dataset and the validation set, respectively. It can be observed that the difference in

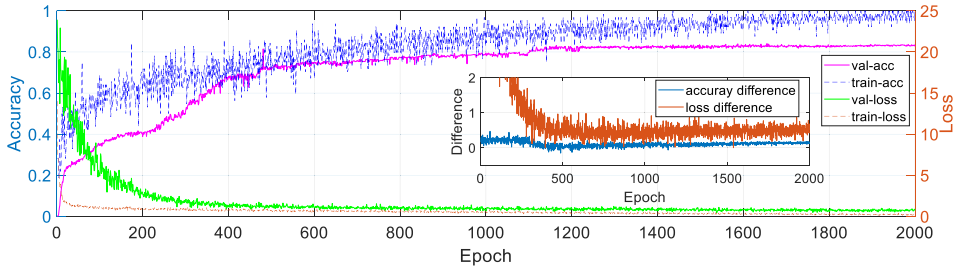

Fig. 5. Performance assessment of model proposed during the training process.

Table 7
Classification accuracy of the proposed model with different data augmentation methods.

Defect types	Initial dataset without augmentation	Conventional method based augmentation	GAN based augmentation	Proposed augmentation solution
Defect-free	0.52	0.59	0.72	0.84
Micro-crack	0.47	0.58	0.74	0.82
Finger-interruption	0.42	0.52	0.71	0.81
Break	0.44	0.56	0.74	0.83

Table 8
Performance comparison with other models using the combined data augmentation method.

	Transfer learning			VGG16		
	MobileNet	ResNet50	InceptionV3	VGG16	VGG16	Proposed solution
Trainable params	2,030,004	2,053,004	2,053,004	25,093,004	25,093,004 39,807,692	12,930,396
Total params	5,258,868	25,640,716	23,855,788	39,807,692	39,807,692	12,930,396
Accuracy	**0.42**	**0.67**	**0.64**	**0.66**	**0.82**	**0.83**

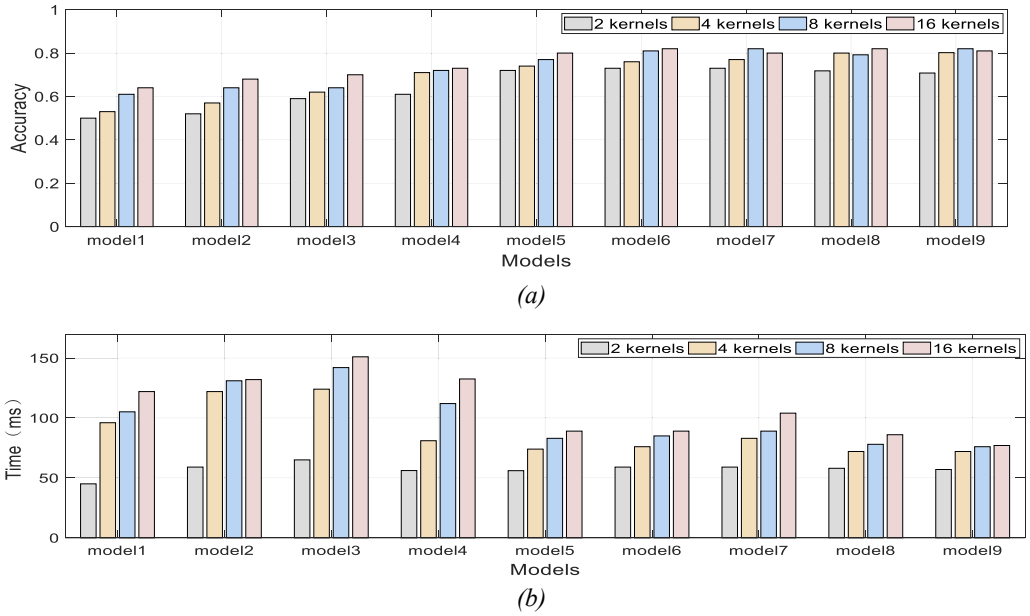

Fig. 6. Experimental evaluation of CNN model with different parameter settings: (a) mean accuracy of defect detection; and (b) mean time consumption of defect detection (model1–model9 represent (c1), (c1c2), (c1c2c3), (c1c2p1), (c1c2p1c3c4p2), (c1c2p1c3c4c5p2), (c1c2p1c3c4p2c5c6p3), (c1p1c2p2), (c1p1c2p2c3p3), respectively).

classification accuracy of the training dataset and validation dataset becomes small along with the increase of the number of epochs, and then becomes significant after 1400 epochs. This is mainly due to the over-fitting problem. However, such an impact on the proposed model is considered not significant as the accuracy performance of the validation dataset remains stable after a sufficient number of epochs.

4.3. Case studies

This section presents and compares the experimental results of the proposed solution against the existing machine learning solutions. The impact of different data augmentation methods, different parameter settings in the CNN based classification model are evaluated and analyzed.

4.3.1. Case 1: Performance comparison with different data augmentation methods

In this case, there are four types of dataset, which are all used for evaluation by training the CNN based model proposed in Section 3.1: the initial dataset, the dataset generated by the traditional image processing method (Section 2.1), the dataset generated by GAN (Section 2.2), the dataset generated by the combined method (Section 2.3). The number of training, validation and test dataset in the other experiments are set as Table 6.

The corresponding accuracy of classification is presented in the Table 7. The value in Table 7 confirms that the accuracy can be further improved by the data augmentation methods and the difference between the three different methods also proves the analysis in Section 2.3. It can be observed that the combined method has the best performance.

4.3.2. Case 2: Performance comparison with the existing methods

In this case, the performance of the proposed CNN based method is compared with the VGG16 (Simonyan and Zisserman, 2015) trained on the EL dataset, as well as a number of transfer learning models, i.e. ResNet50 (He et al., 2016), Inception V3 (Szegedy et al., 2015), and MobileNet (Howard et al., 2017), that are trained based on the ImageNet (Deng et al., 2009). The last layer of these models is replaced by a Global average pooling and two fully connected layers. Other layers are frozen in the models and weights in these frozen layers remain unchanged during the training process.

In this work, two evaluation metrics are particularly considered and analyzed: (1) the classification accuracy for different module defect conditions; and (2) the number of model parameters that could reflect the model complexity. Table 8 presents the details of the number of parameters and the defect classification accuracy for all evaluated models.

It can be observed from Table 8 that, the transfer learning of MobileNet has the minimum number of model parameters, but the defect classification accuracy is significantly lower than the proposed solution. The performance of other transfer learning solutions, i.e., ResNet50, InceptionV3 and VGG16, outperforms the MobileNet in the learning process. The VGG16 that trained on the EL dataset provides equivalent detection performance compared with the proposed solution but with more model parameters and hence higher computational complexity. Overall, the proposed solution outperforms the existing machine learning models in the defect classification of EL images and provides the best trade-off between the classification performance and the model computational complexity.

4.3.3. Case 3: Performance comparison with different parameters selection

In this case, different parameters of CNN are used to evaluate the performance of the model. Considering the selection of different parameters can create numerous different models, only the representative parameter changes are selected in this case.

Impact of the CNN depth: the depth of network can significantly affect the performance of CNN based model, and hence the performance of CNN models with different network depths is assessed. Here, there are in total 9 different CNN structures, i.e. (c1), (c1c2), (c1c2c3), (c1c2p1), (c1c2p1c3c4p2), (c1c2p1c3c4c5p2), (c1c2p1c3c4p2c5c6p3), (c1p1c2p2), (c1p1c2p2c3p3), are extensively evaluated through a comparative study as represented by model1–model9, respectively, in Fig. 6, where (c*i*) and (p*i*) are the i^{th} convolution layers and the i^{th} pooling layer, respectively. Fig. 6 (a) illustrates that the accuracy of the model without the pooling layer is lower than others and when the structure becomes (c1c2p1c3c4c5p2), more convolution layers cannot provide better performance. The result in Fig. 6(b) indicates that adding a more convolution layer can increase the computation complexity and time.

Impact of stochastic pooling: the addition of a pooling layer can obviously decrease the computation complexity as it can decrease the size of the feature map and accelerate the calculation process, as confirmed in Fig. 6(b). The time for defect classification of the model with the pooling layer is much less than those without pooling layers. Here, different combinations of convolution and pooling layer $(c_np)_m$ are evaluated through a comparative study. Here, *c* and *p* represent the convolution layer and pooling layer, respectively, and *n* and *m* are the number of *c* and c_np units, respectively. The performance is evaluated with mean accuracy and the computation time for the same dataset, and the numerical results are shown in Fig. 6. As observed in Fig. 6(a), the accuracy of the model with the pooling layer is obviously higher than those models with only convolution layers.

In addition, it can be observed from Fig. 6(a) that the combination of $(c1c2p1)_n$ and $(c1c2c3p1)_n$ can provide improved performance than the structure of either $(c1c2)_n$ or $(c1c2c3)_n$, and the structure of (c1c2p1c3c4c5p2) outperforms other combinations.

Impact of number of kernels: The results presented in Fig. 6(a) clearly indicate that increasing the number of kernel from 2 to 16 can improve the performance, in particular in (c1), (c1c2), (c1c2c3), (c1c2p1), because the more kernels for convolution can improve the ability to extract the deep feature. However, the improvement effect in deeper network, e.g., (c1c2p1c3c4p2), (c1c2p1c3c4c5p2), (c1c2p1c3c4p2c5c6p3), has diminished. It is can be concluded that when the number of the kernel is beyond 8 or 16, the accuracy can decrease. The result indicates that CNN with 8 or 16 of kernels is sufficient for defect classification of EL images, and an additional number of kernels will increase the computational complexity, as confirmed in Fig. 6(b).

5. Conclusions and future work

This paper proposed a framework for the application of deep learning in the automatic classification of defective PV modules in EL images with a limited sample size. First, effective data augmentation is presented, which can generate high-resolution EL images for deep learning-based methods. Then a CNN based defect classification model is proposed. The classification model is trained with the dataset generated by the combined data augmentation technique. High accuracy is obtained for detecting the defect-free, micro-crack, finger-interruption and break in EL images. The performances of the model are evaluated from multi-perspective. The numerical experimental results also show the efficiency of the model.

For future work, a number of research directions are considered worth further research and development effort. Firstly, various forms of module defects need to be considered and in the performance evaluation of the proposed automatic detection approach using in the EL images. Further investigations and improvements are required for the proposed solution, e.g., efficiently detect multiple coexisting defects in the PV modules using EL images. In addition, the recent technological advances enable the EL sensors and high-performance computing hardware to be mounted in the Unmanned Aerial Vehicles (UAV) to carry out the asset inspection of large-scale PV farms. The lightweight

and efficient machine learning-based models are required to be deployed for online module defect detection during the system inspection.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported in part by the National Key Research and Development Program of China (2018YFB1500900), Natural Science Foundation of China (51777183) and the Major Scientific Project of Zhejiang Lab (No. 2018FD0ZX01).

References

Akram, M.W., Li, G., Jin, Y., et al., 2019. CNN based automatic detection of photovoltaic cell defects in electroluminescence images. Energy 189, 116319.
Alec, R., Luke, M., Soumith, C., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. Comput. Sci.
Alwani, M., Chen, H., Ferdman, M., Milder, P., 2016. Fused-layer CNN accelerators. In: 2016 49th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), Taipei, pp. 1–12.
Arjovsky, M., Bottou, L., 2017. Towards principled methods for training generative adversarial networks. ICLR.
Birkmire, R.W., 2001. Compound polycrystalline solar cells: Recent progress and Y2K perspective. Sol. Energy Mater. Sol. Cells 65 (1), 17–28.
Bouzidi, K., Chegaar, M., Bouhemadou, A., 2007. Solar cells parameters evaluation considering the series and shunt resistance. Sol. Energy Mater. Sol. Cells 91 (18), 1647–1651.
Breitenstein, O., Bauer, J., Bothe, K., Hinken, D., Müller, J., Kwapil, W., Schubert, M.C., Warta, W., 2011. Can luminescence imaging replace lock-in thermography on solar cells? IEEE J. Photovolt. 1, 159–167. https://doi.org/10.1109/JPHOTOV.2011.2169394.
Buerhop-Lutz, C., Deitsch, S., Maier, A., Gallwitz, F., Hauch, J., Camus, C., Brabec, C.J., 2018. A benchmark for visual identification of defective solar cells in electroluminescence imagery. 35th European PV Solar Energy Conference and Exhibition. To appear.
Burrows, K., Fthenakis, V., 2015. Glass needs for a growing photovoltaics industry. Sol. Energy Mater. Sol. Cells 132, 455–459.
Carson, J.A., 2008. Solar Cell Research Progress. Nova Publishers, Hauppauge, NY, USA.
Chegaar, M., Azzouzi, G., Mialhe, P., 2006. Simple parameter extraction method for illuminated solar cells. Solid-State Electron. 50 (7/8), 1234–1237.
Chen, Y., et al., 2017. Investigation of correlation between field performance and indoor acceleration measurements of potential induced degradation (PID) for c-Si PV modules. In: 33rd Eur. Photovolt. Solar Energy Conf. Exhibition, pp. 1408–1413.
Chen, H., Zhao, H., Han, D., et al., 2019. Accurate and robust crack detection using steerable evidence filtering in electroluminescence images of solar cells. Opt. Lasers Eng. 118, 22–33.
Chen, H., Zhao, H., Han, D., et al., 2020. Structure-aware-based crack defect detection for multicrystalline solar cells. Measurement 151, 107170.
Coello, J., 2011. Introducing electroluminescence technique in the quality control of large PV plants. In: Proc. 26th Eur. Photovolt. Sol. Energy Conf. Exhib. pp. 3469–3472.
Deitsch, Sergiu, Christlein, Vincent, Berger, Stephan, Buerhop-Lutz, Claudia, Maier, Andreas, Gallwitz, Florian, Riess, Christian, 2018. Automatic classification of defective photovoltaic module cells in electroluminescence images. Sol. Energy.
Demirci, M.Y., Beşli, N., Gümüşçü, A., 2019. Defective PV cell detection using deep transfer learning and EL imaging. In: Proceedings Book, p. 311.
Deng, J., Dong, W., Socher, R., et al., 2009. Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. IEEE, pp. 248–255.
Dhimish, M., Holmes, V., 2019. Solar cells micro crack detection technique using state-of-the-art electroluminescence imaging. J. Sci.: Adv. Mater. Dev. 4 (4), 499–508.
Frankfurt School-UNEP Centre/BNEF, 2019. Global trends in renewable energy investment 2019. Available: http://www.fs-unep-centre.org.
Fuyuki, Takashi, Kondo, Hayato, Kaji, Yasue, Ogane, Akiyoshi, Takahashi, Yu, 2007. Analytic findings in the electroluminescence characterization of crystalline silicon solar cells. J. Appl. Phys. 101 (2), 023711. https://doi.org/10.1063/1.2431075.
Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al., 2014. Generative adversarial nets. In: Advances in Neural Information Processing Systems (NIPS), pp. 2672–2680.
Gulrajani, I., Ahmed, F., Arjovsky, M., et al., 2017. Improved Training of Wasserstein GANs. Available at https://arxiv.org/abs/1704.00028.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778.
Howard, A.G., Zhu, M., Chen, B., et al., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861.
IEA, 2019. Renewables 2019. Paris https://www.iea.org/reports/renewables-2019.
International Energy Agency, 2018. Renewables 2018, Market analysis and forecast from 2018 to 2023. Available at. https://www.iea.org/renewables2018/.
Jiang, X., Pang, Y., Li, X., et al., 2018. Deep neural networks with Elastic Rectified Linear Units for object recognition. Neurocomputing 275, 1132–1139.
Kang, G.-H., et al., 2015. Prediction of the potential induced degradation of photovoltaic modules based on the leakage current flowing through glass laminated with ethylene-vinly acetate. J. Sol. Energy Eng. 137, 041001–041006.
Köntges, M., Kurtz, S., Packard, C., Jahn, U., Berger, K., Kato, K., Friesen, T., Liu, H., Van Iseghem, M., 2014. Review of Failures of Photovoltaic Modules. Technical Report. International Energy Agency.
Krizhevsky, A., Sutskever, I., Hinton, G., 2012. ImageNet classification with deep convolutional neural networks. International Conference on Neural Information Processing Systems 60, 1097–1105.
Liu, W., Wen, Y., Yu, Z., et al., 2016. Large-Margin Softmax Loss for Convolutional Neural Networks. In: Proceedings of The 33rd International Conference on Machine Learning, pp. 507–516.
Luo, W., et al., 2016. Potential-induced degradation in photovoltaic modules: a critical review. Energy Environ. Sci. 10, 4–7.
Lupangu, C., Bansal, R.C., 2017. A review of technical issues on the development of solar photovoltaic systems. Renew. Sustain. Energy Rev. 73, 950–965.
Martin, A., Chintala, S., Bottou, L., 2017. Wasserstein GAN, 2017. Available at https://arxiv.org/abs/1701.07875.
Mayr, M., Hoffmann, M., Maier, A., Christlein, V., 2019. Weakly supervised segmentation of cracks on solar cells using normalized Lp norm. In: 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, pp. 1885–1889. https://doi.org/10.1109/ICIP.2019.8803116.
Parretta, A., Bombace, M., Graditi, G., Schioppo, R., 2005. Optical degradation of long-term, field-aged c-Si photovoltaic modules. Sol. Energy Mater. Sol. Cells 86 (3), 349–364.
Pazheri, F.R., Othman, M.F., Malik, N.H., 2014. A review on global renewable electricity scenario. Renew. Sustain. Energy Rev. 31, 835–845.
Qian, X., Zhang, H., Yang, C., et al., 2018. Micro-cracks detection of multicrystalline solar cell surface based on self-learning features and low-rank matrix recovery. Sens. Rev. 38 (3), 360–368.
Rosch, R., Krebs, F.C., Tanenbaum, D.M., Hoppe, H., 2012. Quality control of roll-to-roll processed polymer solar modules by complementary imaging methods. Sol. Energy Mater. Sol. Cells 97, 176–180.
Sayigh, A., 2015. Renewable Energy in the Service of Mankind Vol II: Selected Topics from the World Renewable Energy Congress WREC 2014. Springer, New York, NY, USA.
Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations.
Sovetkin, E., Steland, A., 2018. Automatic processing and solar cell detection in photovoltaic electroluminescence images. Integrated Comput.-Aided Eng.
Spataru, S., Hacke, P., Sera, D., 2016. Automatic detection and evaluation of solar cell micro-cracks in electroluminescence images using matched filters. In: 2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC), Portland, OR, pp. 1602–1607.
Szegedy, C., Liu, W., Jia, Y., et al., 2015. Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9.
Tsai, D.-M., Wu, S.-C., Chiu, W.-Y., 2013. Defect detection in solar modules using ICA basis images. IEEE Trans. Ind. Inf. 9, 122–131. https://doi.org/10.1109/TII.2012.2209663.
Yongqing, W., Haipeng, Z., Dengyuan, S., Ailing, C., 2013. Research to the typical defects of crystalline silicon photovoltaic cells based on EL images. Int. J. Energy Sci. 3 (3), 200–204.