

# Defect Detection of Photovoltaic Modules Based on Convolutional Neural Network

Mingjian Sun<sup>(✉)</sup>, Shengmiao Lv, Xue Zhao, Ruya Li, Wenhan Zhang,  
and Xiao Zhang

Department of Control Science and Engineering,  
Harbin Institute of Technology at Weihai,  
Weihai 264209, Shandong, People's Republic of China  
sunmingjian@hit.edu.cn

**Abstract.** Deep learning is employed to detect defects in photovoltaic (PV) modules in the thesis. Firstly, the thesis introduces related concepts of cracks. Then a convolutional neural network with seven layers is constructed to classify the defective battery panels. Finally, the accuracy of the validation set is 98.35%. Besides, the thesis introduces a method in which a single battery cell can be extracted from the Electro Luminescence (EL) image of the PV module. This method is very suitable for automatic inspection of photovoltaic power plants.

**Keywords:** Convolutional Neural Network · PV module cracks  
Defect detection · Deep learning

## 1 Introduction

Photovoltaic power generation has become the most widely used way of generating new energy. In December 2016, National Energy Administration of China announced that by the end of 2016, cumulative installed capacity of photovoltaic power of China had reached 77.42 million kilowatts. In accordance with national requirements, China's photovoltaic power capacity will reach more than 150 million kilowatts in 2020 (see [1]). The core component of the whole photovoltaic power plant is the solar panel. The inevitable defects in the production and installation process will affect the efficiency of the plant. Thus, it is necessary to carry out defect detection for solar panels. The existing detection methods which are relatively mature in application are Infrared Thermal Imaging (ITI) and Electro Luminescence (EL). Infrared thermal imaging is suitable for a wide range of detection, but generally this approach only detects hot spot defects. The Electroluminescence method is suitable for detecting defects in a single PV module. Compared with infrared thermal imaging, it can show the details of the defects more clearly. It is generally used for the detection of the hidden cracks in the single module. At present, for large-scale photovoltaic power plants, manual sampling method is generally adopted, which costs a lot of labor and time. In terms of the current method of manual sampling, this paper proposes a kind

of automatic detection method based on deep learning, which can realize the automatic detection and classification of the hidden cracks of PV modules.

Convolutional Neural Network is a classic deep learning framework inspired by the biological perception of natural visual perception. Convolutional neural networks, characterized by translational invariance, shared values and pooling, are effective in reducing network parameters, which renders it outstanding performance in many areas like image processing, video and voice (see [4]). In this paper, a CNN with seven layers is established to identify and classify the hidden flaws. The final classification accuracy reaches 98.35%.

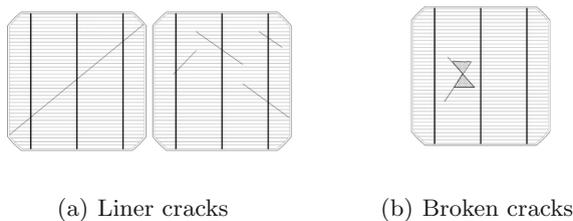
The thesis introduces related concepts of cracks and the hazards and classification of hidden cracks in the Sect. 2. The theory of CNN, including the establishment of data set, the structure of neural network, the algorithms and so on, is introduced and the obtained results are shown in the Sect. 3. Section 4 describes how to extract a single battery cell from a single PV module.

## 2 Cracks

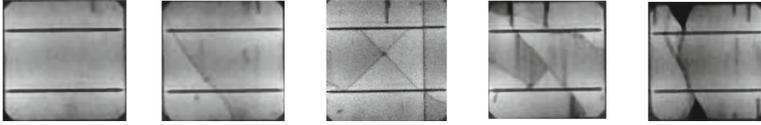
When the battery cell (component) is subjected to greater mechanical or thermal stress, the invisible crack probably comes into being, which is difficult to detect (see [10]). Different from hot spots, cracks only lead to battery disconnection, thus affecting the power output. Different types of cracks have different effects on the panels. As the hidden crack is difficult to directly observe with eyes, EL test is necessary for observation.

### 2.1 The Hazards and Classification of Cracks

The current flow path in the battery is that the collected current is transmitted to the main grid line by the fine grid line and is led out through the bus bar and the junction box. The current of the battery chip is proportional to the area of the power generation. If part of the current can not be transmitted to the main grid line due to cracks, the power output of the PV module will be affected. Thus, the main hazard of crack is forming failure area and affecting the output power (see [11, 12]). Figure 1(a) shows that the hidden crack runs through the battery unit, but does not form a failure area, so the impact on the power output



**Fig. 1.** Different cracks.



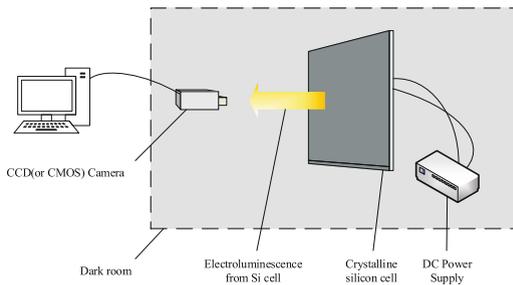
**Fig. 2.** From left to right, the types of cracks are normal, linear, cross, flaky and broken. The impact on output power is from weak to strong.

is less. Figure 1(b) shows that the battery cell has a failure area, therefore there is a worse impact on the power of the PV module.

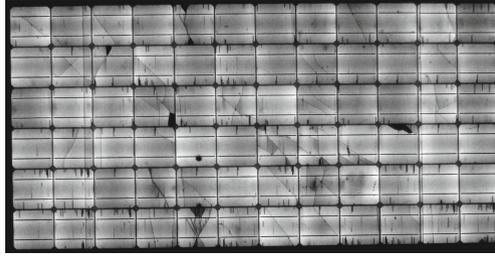
Therefore, according to the magnitude of hazards, the cracks are divided into five categories, respectively: normal, linear, cross, flaky and broken cracks, based on the impact on power output from weak to strong, as is shown in Fig. 2.

## 2.2 Electroluminescence Mechanisms

When the crystalline silicon cell is applied with a forward bias voltage, the carrier traverses the PN junction so that the carrier concentration exceeds the thermal equilibrium value to form an excess carrier. Excessive carrier recombines and the energy is released in the form of heat and light (photon). In the photon emission, the electrical energy is transformed into light, which is called injection electroluminescence (see [9]). Emission spectrum is mainly concentrated between 1000 nm–1300 nm. The brightness of electroluminescence is proportional to the total number of minority unbalanced carrier, the minority-carriers diffusion length and the current density. The minority diffusion length is lower where there is a defect, so the image is dark relatively. Thus, the defect can be judged from the light and shade. Figure 3 shows the EL mechanisms of test equipment. The camera used for the experiment is OPT-M311, the main chip of which is Sony EXPEED4, 24 million pixels. Figure 4 shows the image acquired by the EL detector.



**Fig. 3.** Electroluminescence mechanisms.



**Fig. 4.** Images acquired by EL detector. Defects can be judged from the light and shade.

### 3 Convolutional Neural Network

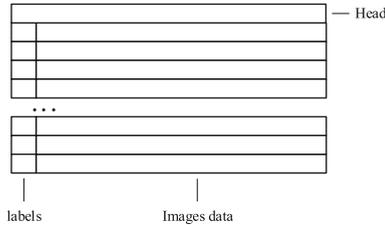
#### 3.1 Image Data Set

The dataset includes the EL images acquired during the practical process. The EL image of each component is divided into images of single battery cell and processed to grayscale images with a size of  $100 \times 100$ . The original data can not be directly input to the neural network. Standardization is required and the method is as follows:

$$y = \frac{(x - mean)}{stddev} \tag{1}$$

where  $x, y$  are the pixel values at any point of the original image before and after the transformation.  $mean$  is the mean of the image.  $stddev$  is the standard deviation of the image.

It should be noted that the input image can not be uniform, otherwise there will be a division by 0 error. The entire data set storage structure is shown in Fig. 5.



**Fig. 5.** The storage structure of data set

The establishment of the data set refers to the format of the MNIST database of handwritten digits (see [14]), consisting of file headers, labels, pictures. The headers record the number and the ranks of images. In this way, the labels and images are combined together. Although the establishment process is slightly cumbersome compared to the data sets where labels and pictures are separated,

this way has its own advantages. Firstly, data sets can be read into the memory block by block, avoiding the lack of memory when the data set is too large and avoiding the waste of time by reading pictures one by one. Secondly, it can facilitate the expansion of data. Comparatively speaking, data sets where labels and pictures are separated can only be read into memory all at once, and it is difficult to expand the data set.

### 3.2 Data Augmentation

Small data sets easily lead to over-fitting, while large data sets are difficult to obtain. Regarding this, data augmentation is quite essential (see [3]). The main methods include rotating the images, adjusting the brightness, conducting horizontal transformation and blurring the images. Rotating the images once every  $5^\circ$  from  $-10^\circ \sim 10^\circ$  can enhance the effect from the slight tilt of the captured image. In addition, it is necessary to rotate  $90^\circ$ ,  $180^\circ$  and  $270^\circ$  in turn. This will reduce the influence of the main grid lines on the classification accuracy, as the main line is easily recognized as a defect. Brightness and blurring are adopted to simulate the common situations of capturing images with cameras in order to enhance the generalization ability of neural network. Different methods of data augmentation are shown in Fig. 6. Finally, the entire data set has a total of 6120 pictures, of which 5120 images are used as training set and 1000 images are for validation.

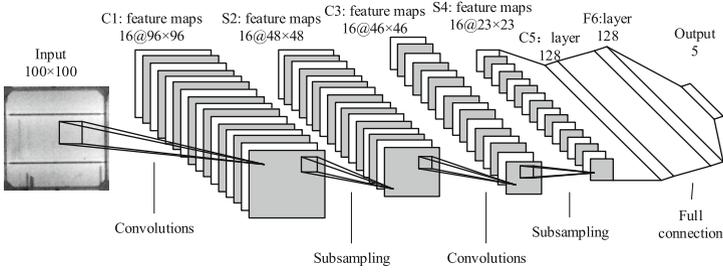


**Fig. 6.** The methods of data augmentation. From left to right: original image, rotating  $90^\circ$ , increasing brightness, mirror transformation, rotating  $5^\circ$ .

### 3.3 Network Architecture

The CNN constructed in this paper is a multi-layer structure referring to [5], including two convolutional layers (C1, C3), two pooling layers (S2, S4), and a fully connected network. The overall network structure is shown in Fig. 7. As is shown, this network is relatively simple. It is mainly because the features of hidden cracks are obvious and easy to identify. This shallow network can achieve better results and effectively reduces the amount of calculation and the difficulty of training, thus this network is of engineering practicability.

The input image is a  $100 \times 100$  grayscale image. After convolving with sixteen  $5 \times 5$  kernels, sixteen  $96 \times 96$  feature maps (C1) are obtained. After  $2 \times 2$  max-pooling, sixteen  $48 \times 48$  maps (S2) are acquired. Next, sixteen  $46 \times 46$  feature maps (C3) are obtained with sixteen  $3 \times 3$  kernels and then with max-pooling, layer S4 is produced. Finally, followed is a three-layer structure of fully connected network with 128 neurons in C5 layer and F6 layer. The last layer is the output layer with five neurons meaning five categories.



**Fig. 7.** The structure of CNN

Apart from the last layer using the Sigmoid activation function, the network uses the ReLU activation function (see [4]). The advantages are:

1. The essence of ReLU is a piecewise linear model. Forward calculation is very simple without calculation like exponent calculation.
2. ReLU is easy to calculate partial derivatives in back propagation.
3. ReLU is easy to train. The derivative will not tend to zero like Sigmoid activation function.
4. When input is less than 0, ReLU outputs 0. Thus many neurons output is 0, making the network become sparse, which can reduce the over-fitting phenomenon.

As the total output value of Softmax is 1, which means that the growing probability of one term will inevitably lead to the reduction of others, that is, output results of softmax are exclusive, the last layer does not use the Softmax function. However, a variety of defects may exist on the same cell of the battery. Softmax may lead to that the probabilities of different defects are all very small. Assuming that there are three defects on the battery chip, the output probabilities of defects may be 0.3, 0.3 and 0.4. Then it is not easy to judge what flaws are.

### 3.4 Training and Results of Neural Network

As mentioned earlier, increasing the training samples is one way to reduce over-fitting. Another is L2 normalization (weight decay). The idea of L2 normalization is to add an additional item to the loss function (see [13]). The regularization term shown in Eq. 2 is added to the cross entropy loss function which is used in this thesis.

$$C = -\frac{1}{n} \sum_{x_j} [y_j \ln a_j^I + (1 - y_j) \ln(1 - a_j^I)] + \frac{\lambda}{2n} \sum_w w^2 \quad (2)$$

where  $x$  is the input sample,  $a$  is the actual output vector of the neural network,  $y$  is the desired output vector for the neural network,  $w$  is the weight. The first term in Eq. 2 is just the usual expression for the cross-entropy. The second term, namely the sum of the squares of all the weights in the network is added to the cross-entropy. This is scaled by a factor  $\frac{\lambda}{2n}$ , where  $\lambda > 0$  is known as the regularization parameter.

In the practical training process, the dropout method is also applied [3]. The dropout rate is 50%, which means that the connection layer neuron output is set to 0 randomly. In this way, the corresponding weight will not be updated, which can help reduce over-fitting. The training method applies the Nesterov gradient acceleration (NAG) method, which is slightly different from the momentum update and stochastic gradient descent and has become more popular now [2]. Nesterov can guarantee a stronger theoretical convergence for convex function and the practical performance is better than momentum and stochastic gradient descent. The core idea of Nesterov is to accelerate in the same gradient direction and decelerate in the changing gradient. The momentum update formula is Eq. 3.

$$\Delta V_t = \rho \Delta V_{t-1} - \eta [\nabla C (V_{t-1})]^T \quad (3)$$

The momentum update includes two parts: one is the updated value  $\rho \Delta V_{t-1}$ , which has been calculated at last moment and is known at this moment; the other is calculated gradient  $\eta (\nabla C)^T$  based on the current position. Nesterov accelerated gradient points out: since it is known that the update at this time will go  $\rho \Delta V_{t-1}$ , then go  $\rho \Delta V_{t-1}$  first and correct according to the gradient there. Therefore, the gradient calculated in this method is not on a basis of old position, but is forward-looking.

$$\Delta V_t = \rho \Delta V_{t-1} - \eta [\nabla C (V_{t-1} - \rho \Delta V_{t-1})]^T \quad (4)$$

The NAG update formula is shown in Eq. 4. In Eq. 4, the learning rate is set to 0.005, the momentum factor is set to 0.9, the batch size is set to 64, with a total of 2000 steps, that is 25 epochs.

Figure 8 shows the relationship between the accuracy of the classification and the number of training steps. The solid line indicates the accuracy rate of the training set. The test is to extract 1000 images randomly from the training set

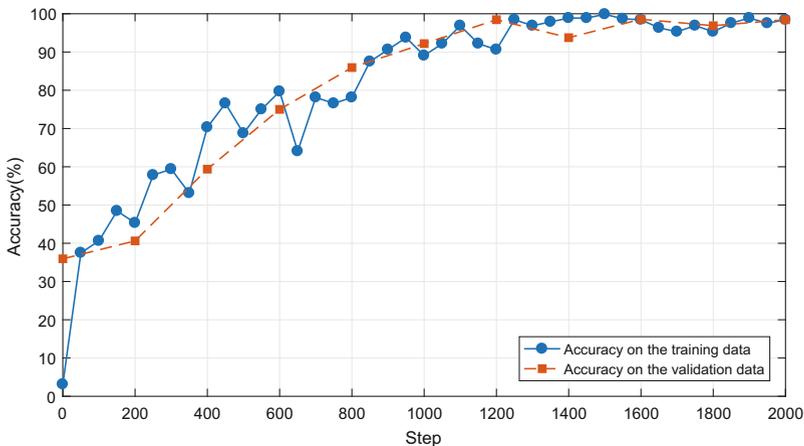
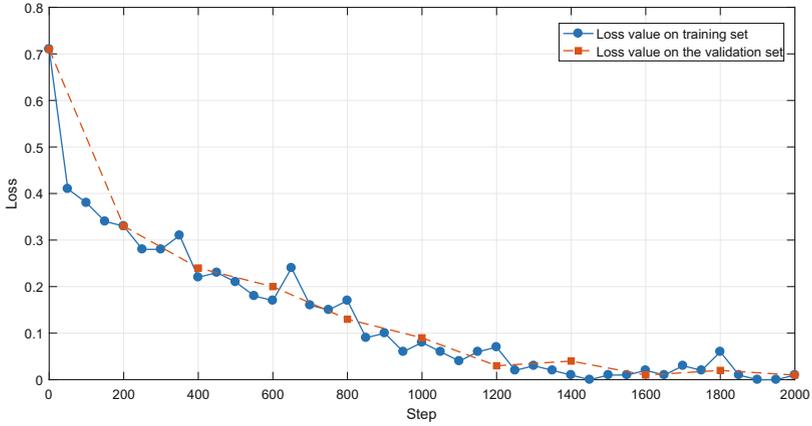
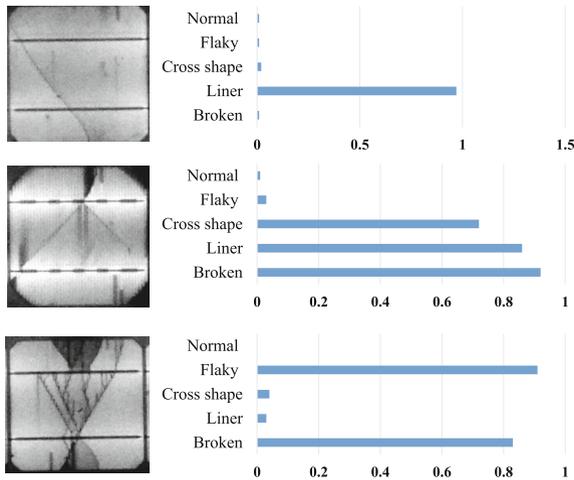


Fig. 8. Accuracy on the training and validation data set.



**Fig. 9.** The loss function is changing with the number of steps on the training and validation data set. The value of the loss function is small, indicating that the neural network has achieved better performance.

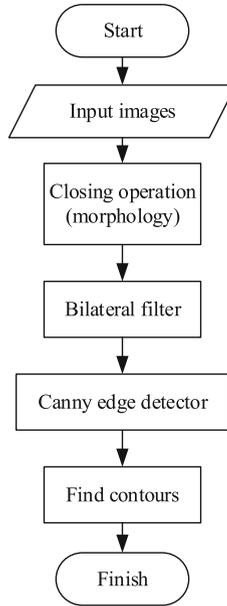


**Fig. 10.** Testing a single picture. The horizontal coordinate represents the probabilities predicted by the neural network and vertical coordinate represents the types of cracks.

to test the accuracy every 50 training steps. At Step 2000, the accuracy of the training set is 98.40%. The dotted line indicates the accuracy of the validation set. During the training process, there is validation on validation set every 200 steps. The final accuracy is 98.35%. As is shown, the network have achieved relatively good results on the test set and the validation set. The results indicate that the network is reasonable and is of strong generalization ability without over-fitting. This point can be seen from Fig. 9. The loss function finally reached a very small value of 0.2. Figure 10 shows the probabilities of different defects when testing a single image.

## 4 Acquirement of the Battery Unit EL Images

As is shown in Fig. 3, the first thing is to obtain crack images from the camera, and then use the neural network to identify the crack. The camera captures the EL image of the entire PV module, but the neural network requires an image of a single cell. So it is necessary to pre-process the acquired EL images. The processing steps are shown in Fig. 11.



**Fig. 11.** The process of images processing

First, the morphological closed operation is used to fill the small voids in the foreground, especially reducing the effect of the two main grid lines on each cell. Otherwise, the two main grid lines will cause negative effects on border extraction. Second, use the bilateral filter algorithm to filter the pictures (see [8]). Bilateral filtering algorithm is a nonlinear filtering method, the advantage of which is to retain the edge information when filtering. This is favorable for the edge extraction. Next, use Canny edge detection to extract the edges (see [6]). Finally, adopt the method in literature [10] to extract the boundaries of each cell. The whole result is shown in the Figs. 12 and 13.



modules tilt in the range of  $-20^\circ \sim 20^\circ$ . Thus, the system is very suitable for auto-inspecting the photovoltaic power plant, such as using a unmanned aerial vehicle (UAV) equipped with cameras to inspect power station.

## References

1. China National Energy Administration: 13th five year plan for solar energy development. *Solar Energy* **12**, 5–14 (2016)
2. Bengio, Y., Boulanger-Lewandowski, N., Pascanu, R.: Advances in optimizing recurrent networks. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8624–8628. IEEE (2013)
3. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*, pp. 1097–1105 (2012)
4. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**(7553), 436–444 (2015)
5. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. *Proc. IEEE* **86**(11), 2278–2324 (1998)
6. Rong, W., Li, Z., Zhang, W., Sun, L.: An improved canny edge detection algorithm. In: 2014 IEEE International Conference on Mechatronics and Automation (ICMA), pp. 577–582. IEEE (2014)
7. Suzuki, S., et al.: Topological structural analysis of digitized binary images by border following. *Comput. Vis. Graph. Image Process.* **30**(1), 32–46 (1985)
8. Zhang, B., Allebach, J.P.: Adaptive bilateral filter for sharpness enhancement and noise removal. *IEEE Trans. Image Process.* **17**(5), 664–678 (2008)
9. Fuyuki, T., Kitiyanan, A.: Photographic diagnosis of crystalline silicon solar cells utilizing electroluminescence. *Appl. Phys. A* **96**, 189–196 (2009)
10. Wu, Z., Huang, H.: Research on the effects of PV modules' transformation on the performance. *Mech. Eng. Autom.* **4**, 107–109 (2011)
11. Xu, Z., Wang, H., et al.: Research progress in crack features of PV modules: part 1. *Solar Energy* **10**, 47–51 (2015)
12. Kajari-Schröder, S., Kunze, I., Eitner, U., Köntges, M.: Spatial and orientational distribution of cracks in crystalline photovoltaic modules generated by mechanical load tests. *Solar Energy Mater. Solar Cells* **95**, 3054–3059 (2011)
13. Krogh, A., Hertz, J.A.: A simple weight decay can improve generalization. In: *NIPS*, vol. 4, pp. 950–957 (1991)
14. MNIST database of handwritten digits. <http://yann.lecun.com/exdb/mnist/>