Automatic Detection of Defective Solar Cells in Electroluminescence Images via Global Similarity and Concatenated Saliency Guided Network

Jinxia Zhang, Yu Shen, Jiacheng Jiang, Shixiong Fang, Liping Chen, Tingting Yan, Zuoyong Li, Kanjian Zhang, Haikun Wei, Weili Guo

***Abstract*— Electroluminescence imaging becomes a very useful technique to automatically detect defects for solar cells since it can provide high resolution electroluminescence images. However, few methods explicitly consider the visual characteristics of the defects and the noises in solar cells. In this paper, a global pairwise similarity and concatenated saliency guided neural network is proposed by fully considering the observed visual characteristics in electroluminescence solar cell images. The proposed network exploits a global pairwise similarity module and a concatenated saliency module to refine the features extracted by the convolutional neural network. The global pairwise similarity module aims to refine the features of an image pixel by modeling long-range dependencies. The concatenated saliency module is exploited to suppress the background and decouple different salient regions to better represent the features of an image. Extensive experiments based on five different baselines, i.e. VGG16, ResNet56, ResNet50, DenseNet40 and GoogleNet, prove that the proposed method significantly outperforms the baseline models and show that both the global similarity module and the concatenated saliency module can help to detect defective solar cells in electroluminescence images.**

***Index Terms*— Attention, automatic defect detection, electroluminescence image, saliency, solar cell.**

Manuscript received October 18, 2021; revised February 06, 2022; accepted September 18, 2022. This work was supported in part by the National Key R&D Program of China(2018YFB1500800), Guangdong Basic and Applied Basic Research Foundation(2022A1515011435), ZhiShan Scholar Program of Southeast University, the National Natural Science Foundation of China(61906092), Natural Science Foundation of Jiangsu Province of China(BK20190441), the Fundamental Research Funds for the Central Universities, Science and Technology Project of State Grid Corporation of China(SGTJDK00DYJS2000148) and the Big Data Computing Center of Southeast University. (Corresponding authors: Haikun Wei and Weili Guo)

Jinxia Zhang is with Key Laboratory of Measurement and Control of CSE, Ministry of Education, School of Automation, Southeast University, Nanjing, China and with Southeast University Shenzhen Research Institute, Shenzhen, China (email: jinxiazhang@seu.edu.cn).

Yu Shen, Jiacheng Jiang, Shixiong Fang, Kanjian Zhang and Haikun Wei are with Key Laboratory of Measurement and Control of CSE, Ministry of Education, School of Automation, Southeast University, Nanjing, China (email: 230169411@seu.edu.cn; alanjjc1998@163.com; sxfang@seu.edu.cn; kjzhang@seu.edu.cn; hkwei@seu.edu.cn).

Liping Chen and Tingting Yan are with Jiangsu (SunTech) Institute for Photovoltaic Technology, Wuxi, China (email: lpchen@suntech-power.com; minzhouTM@suntech-power.com).

Zuoyong Li is with Fujian Provincial Key Laboratory of Information Processing and Intelligent Control, Minjiang University, Fuzhou, China (email: fzulzytdq@126.com).

Weili Guo is with School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China (email: wlguo@njust.edu.cn).

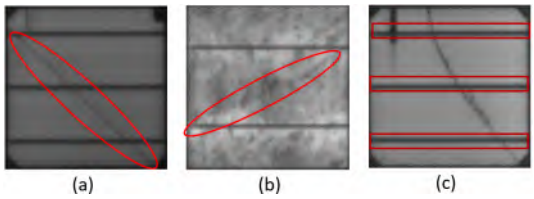

Fig. 1. Examples of defective solar cells. In this figure, (a) and (c) are monocrystalline silicon solar cells, and (b) includes a polycrystalline silicon solar cell.

I. Introduction

ENERGY production from photovoltaic (PV) power stations has substantially increased over the past decades. The statistics data reports that about 362 GWp of crystalline silicon solar modules have been installed worldwide [1]. The energy efficiency of solar modules which determines the performance of a power station is highly correlated with their health condition [2], [3]. In reality, solar modules can be inevitably damaged caused by the dropping during installation or the falling of the tree branches. Moreover, faulty soldering or defective wires during the actual production process can also cause the damage of solar modules.

Electroluminescence (EL) imaging is a non-destructive technique, which can provide high resolution EL images and have a good ability to detect small defects [4]. Thus, EL becomes a very useful and popular modality for automatic defect detection of solar modules [5]. Examples of EL images can be seen in Fig. 1. Fig. 1(a) and Fig. 1(c) are monocrystalline silicon solar cells, and Fig. 1(b) includes a polycrystalline silicon solar cell. The monocrystalline silicon solar cell has a uniform background texture and the polycrystalline silicon solar cell has a complex background texture [6]. Besides, intrinsic crystal grain boundaries and extrinsic defects of micro cracks and breaks have lower intensity. Recent few years, some works have demonstrated that deep learning based

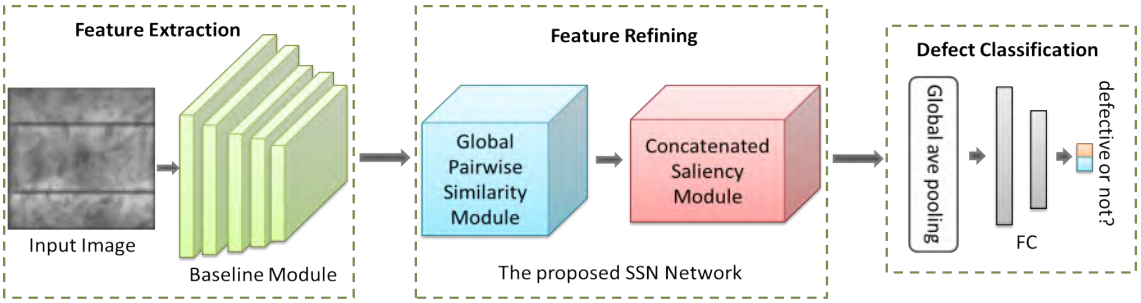

Fig. 2. Overview of the proposed method. For simplicity, the proposed global pairwise Similarity and concatenated Saliency guided Neural network is abbreviated as SSN.

defect detection models can outperform the traditional pattern recognition methods [7], [8] to detect defects in EL images. However, few deep learning based methods explicitly consider the visual characteristics of the defects and the noises in solar cells. In this paper, a global similarity and concatenated saliency guided neural network is proposed which exploits a global pairwise similarity module and a concatenated saliency module, by fully considering the observed characteristics in EL solar cell images.

According to our observation, lots of defects in EL solar cell images are long and thin, shown in the red circles of Fig. 1(a) and Fig. 1(b), which include one of the common defects named micro-crack. In the widely used convolutional neural networks (CNN), the fundamental module is the convolution block, where the discrete convolution operation is computed based on a sliding local patch and a small-sized convolution kernel. Thus, long-range dependencies in images can only be captured when the convolution operations are processed repeatedly for CNN. However, repeating local operations usually make the features of the deep layers blurry and part features of thin objects in the input image may disappear in the deep layers. To solve this problem, a global pairwise similarity module is exploited to refine the features extracted by the CNN. Specifically, the similarity between each image pixel and each representative patch is computed and the features of an image pixel are refined according to the features of all similar patches.

Another observation of the solar module cell is that there are a lot of noises in the EL images. One of the noises is the horizontal gratings, shown in red boxes of Fig. 1(c). The gratings make it harder to identify the true defects. Moreover, the background of the polycrystalline silicon solar module cell (Fig. 1(b)) is complex and contains irregular floccules. To handle this problem, a concatenated saliency module is exploited to decouple and concatenate different salient features to better represent a solar module cell.

Based on the above observations, to improve the feature representation capability of the CNN, a global pairwise similarity and concatenated saliency guided neural network is proposed to automatically detect defective EL solar cell images. More specifically, a global pairwise similarity module is exploited to refine the features of the defects in EL solar cell images, which are usually long and thin. Furthermore, a concatenated saliency module is exploited to effectively suppress the background and extract diverse potential salient features in images. With the help of the global pairwise similarity module and the concatenated saliency module, the proposed network can better represent the features of the defects.

For simplicity, the proposed global pairwise Similarity and concatenated Saliency guided Neural network is abbreviated as SSN in the paper. The proposed SSN network is illustrated in Fig. 2. The basic features of the input image are extracted via a baseline convolutional module, such as VGG. The image features are further refined via the proposed network: Firstly, the image features are refined based on the global pairwise similarity module; Secondly, the image features are refined according to the concatenated saliency module. The refined features help to predict whether the input EL image is defective or not.

A. Contributions

This work aims to automatically detect defective solar cells in EL images by proposing a new method and the contributions of this work are as follows.

Firstly, a global pairwise similarity and concatenated saliency guided neural network is proposed to automatically detect whether an EL solar cell image contains defects or not. This network aims to refine the features extracted by the baseline CNN to better detect defects in EL images. Extensive experiments demonstrate that the proposed network significantly outperforms the baselines.

Secondly, a global pairwise similarity module is exploited to capture long-range dependencies and better represent long and thin defects in solar module cells. Experiments prove that the global pairwise similarity module helps to recall the defective EL images.

Thirdly, a concatenated saliency module is exploited to decouple the features of diverse defects and other potential noises. Experiments show that the concatenated saliency module can help to improve the performance of precision and detect the true defects.

B. Outlines

The remainder of this paper is organized as follows: related works are introduced and discussed in Section II; the proposed network is described in Section III, including the global pairwise similarity module and the concatenated saliency module; Section IV illustrates our testing procedure and experimental comparison based on five different baselines. Finally, the conclusion of the paper is made in Section V.

II. Related work

EL images provide high visual resolution, which are useful to detect finest defects of solar modules. However, EL image can not only exhibits barely visible defects as dark objects, but also shows random dark regions in the background, which makes automatic defect detection in EL images difficult. Visual analysis of EL images requires abundant expert knowledge, which is expensive and time-consuming. Thus, automatic defect detection of solar module cell in EL images becomes more and more attractive and active.

Most existing works employed traditional pattern recognition methods to solve this problem. Tsai et al. propose a Fourier image reconstruction technique to detect defect of solar module cells with EL images [9]. Since the defects usually appear as line-shaped objects in the EL images, Tsai et al. set the frequency components associated with the line-shaped defects to zero and transformed the spectral image back to a spatial image. The defect can be detected by computing the difference between the original image and the Fourier reconstructed image. Dhimish et al. also detect micro-cracks in frequency domain based on the binary and discrete Fourier transform image processing models [10]. Wang et al. classify solar cell defect by computing the absolute EL intensity loss rate based on solar cell EL images and image information [11]. Dhimish et al. calibrate the EL image using a bit-by-bit gridding technique and an OR gate between each bit of the nondefective and cracked solar cells [12]. Tsai et al. propose a defect detection method based on independent component analysis (ICA) for EL images [13]. Su et al. propose a discriminable feature descriptor named as center pixel gradient information to center-symmetric local binary pattern to detect defect in EL images [14]. Deitsch et al. propose a data-driven method to detect defective photovoltaic module cell via a support vector machines classifier (SVM) [7]. In this method, keypoints are first detected and the features are described via SIFT, SURF and KAZE. Finally, the feature descriptors are combined via SVM.

Since 2012, CNN have revolutionized computer vision tasks [15]. Successively deep convolutional neural networks have been proposed for image recognition [16]–[19]. The CNN architecture was first introduced to detect defects for EL images by Deitsch et al [7]. Deitscha et al. employ a CNN model, i.e. VGG [18] to automatically classify the defective photovoltaic module cells in EL images. In this work, the authors also prepare the performance of VGG with the performance of SVM based on hand-crafted features and demonstrate that the CNN is more accurate. Karimi et al. also find CNN to outperform SVM and Random Forest (RF), with SVM performing the second best and RF showing the lowest performance [8].

Since then, a few deep learning based methods are proposed to deal with the problem of the uncertainties of the labels created by experts and the lack of positive samples. To deal with the data uncertainties, Ge et al. propose a hybrid fuzzy convolutional neural network to detect cell defect [20]. To deal with the data scarcity, Shou et al. propose a model based on generative adversarial network (GAN) and auto-encoder (AE) to detect defect for EL images [21]. Tang et al. propose an EL image sample generation method by combining traditional image processing technology and GAN [22]. Akram et al. generate EL images based on data augmentation operations [23].

To deal with the similarity between the defect features and the complex background features in EL images, Su et al. connect the channel-wise attention and the spatial attention to Faster RCNN [24], [25]. This work detects cell defect based on small regions instead of the whole EL image. Thus, the performance of this method relies on the quality of the region proposals. Aside from the complex background, very few deep learning based works explicitly consider the visual characteristics of the cell defects and the noises in the EL images. This work attempts to refine the features extracted by the popular CNN models and propose a global pairwise similarity and concatenated saliency guided neural network by considering the visual characteristics in EL images.

III. The proposed method

According to our observations, the defects in solar module cells are usually long and thin. Also, there are noises in EL images, such as gratings and irregular floccules in polycrystalline solar module cells which may prevent to detect true defects. This work aims to optimize the convolutional neural network and improve the feature representation of defects. To this end, the global similarity and concatenated saliency guided neural network is proposed. Two modules, i.e. the global pairwise similarity module and the concatenated saliency module are exploited in this network and added to the convolutional neural network to automatically identify whether the input image contains defects or not.

A. Global pairwise similarity module

The global pairwise similarity module explores the similarity between a pixel and multiple representative patches in the whole image and updates the features of this pixel according to the weighted sum of the features of all similar patches. Assume the feature tensor generated by a baseline convolution module is $F \in \mathbb{R}^{C\times H\times W}$, where C denotes the number of feature channels, while H and W represent the height and the width of the features, respectively. The feature tensor F is usually large and it would require large amount of memory if the similarities is computed between each pair of images pixels.

To reduce the memory requirement and increase computational efficiency, the features of representative patches P are first computed according to the feature tensor F. An input feature F is divided into $G \times G$ grids as representative patches and the max feature value is computed in each grid as the feature of this grid. Thus, the size of the features for representative patches P is $C \times G \times G$, where $G \leq H$ and $G \leq W$.

To further improve the computational efficiency, an 1×1 convolutional kernel is applied to the feature tensor F and the features of representative patches P to reduce the number of the channels from C to C'. The similarity of a pixel $F_i(1 \leq$

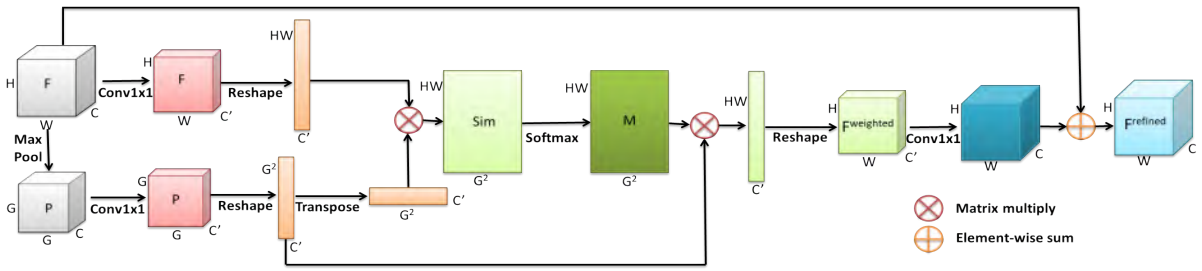

Fig. 3. The detailed architecture with step-by-step operations of the global pairwise similarity module.

$i \leq HW$) and a representative patch $P_j (1 \leq j \leq G^2)$ is then computed according to the reshaped features:

$$Sim(F_i, P_j) = F_i P_j^T \tag{1}$$

After computing the similarity between each image pixel and each representative patch, a similarity matrix $Sim \in \mathbb{R}^{HW \times G^2}$ can be got. Then the weight matrix $M \in \mathbb{R}^{HW \times G^2}$ is computed based on the similarity matrix Sim.

$$M = softmax(Sim) \tag{2}$$

In this equation, the function *softmax* is employed to normalize the weights. The similarity weighted feature $F_i^{weighted}$ is then computed based on the weighted sum of the features of the representative patches, where each weight is $M(F_i, P_j)$.

$$F_i^{weighted} = \sum_{j=1}^{G^2} M(F_i, P_j) P_j \tag{3}$$

An 1×1 convolutional kernel is then applied to $F_i^{weighted}$ to change the number of channels from C' back to C. And the refined feature $F_i^{refined}$ is computed as the sum of the weighted feature and the input feature.

$$F^{refined} = conv1(F^{weighted}) + F \tag{4}$$

The detailed architecture of the global pairwise similarity module with step-by-step operations is shown in Fig. 3. The outputs after each operation and its size are indicated in the corresponding positions. The sizes of the input feature and the output feature after the global pairwise similarity module are the same.

B. Concatenated saliency module

Recently, the field of saliency has gained its popularity since it can be applied to many computer vision tasks, including object recognition and person re-identification [26]–[29]. Inspired of the work [28], concatenated saliency module is exploited to suppress the complex background and decouple diverse defects and potential noises in EL images.

1) Saliency module: For the input feature tensor $F \in \mathbb{R}^{C \times H \times W}$, an 1×1 convolutional layer is first employed to linearly combine different channels and get a concise feature matrix $X \in \mathbb{R}^{H \times W}$. According to our observation, a solar module cell usually contains horizontal gratings, thus a horizontal stripe based average pooling operation is applied to the concise feature map X and generate a compact saliency descriptor vector $V = (v_1, v_2, ..., v_H)^T$, which size is $H \times 1$. The compact saliency descriptor vector V is then normalized

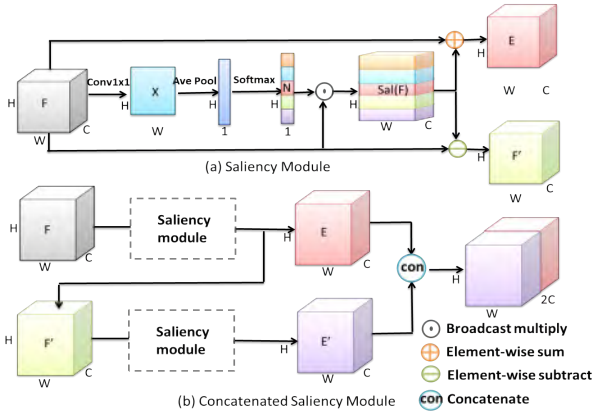

Fig. 4. The detailed architecture with step-by-step operations of saliency module and concatenated saliency module.

with a *softmax* activation function to get a normalized saliency-sensitive weight vector $N = (n_1, n_2, ..., n_H)^T$.

$$n_i = \frac{exp(v_i)}{\sum_1^H exp(v_j)}, i \in [1, H] \tag{5}$$

The saliency tensor is further computed using a broadcast multiplication based on the normalized saliency-sensitive weight vector N and the input feature tensor F. The vector N would be first broadcasted along the channel and the weight dimensions to have the same shape as F and then multiplied with F to get the saliency tensor $Sal(F)$.

$$Sal(F) = N \odot F \tag{6}$$

And the saliency enhanced feature E is then computed by integrating the saliency tensor $Sal(F)$ and the base input feature tensor F.

$$E = Sal(F) + F \tag{7}$$

To further mine other potential salient features and represent diverse image features, the current saliency tensor $Sal(F)$ is suppressed by subtraction to get the input feature for the next saliency module.

$$F' = F - Sal(F) \tag{8}$$

This suppression operation aims to alleviate the effect of the current saliency $Sal(F)$ on other important features and highlight other potential salient information.

2) Concatenated saliency module: The saliency suppressed feature F' is then input to another saliency module and get another saliency enhance feature E' in the same manner as introduced in Sec. III-B.1. And all saliency enhance features are concatenated together to get the final refined features.

TABLE I
THE CONFIGURATIONS OF DIFFERENT BASELINE MODULES.

VGG16	ResNet56	ResNet50	DenseNet40	GoogleNet
Input (300x300x3 EL image)				
[3x3,64] x 2 2x2 max pool	3x3,16	7x7,64 3x3 max pool	3x3,24	3x3,192
[3x3,128] x 2 2x2 max pool	[3x3,16 3x3,16] x 9	[1x1,64 3x3,64 1x1,256] x 3	[3x3,12] x 12 1x1,168 2x2 ave pool	[Inception] x 2 3x3 max pool
[3x3,256] x 3 2x2 max pool	[3x3,32 3x3,32] x 9	[1x1,128 3x3,128 1x1,512] x 4	[3x3,12] x 12 1x1,312 2x2 ave pool	[Inception] x 5 3x3 max pool
[3x3,512] x 3 2x2 max pool	[3x3,64 3x3,64] x 9	[1x1,256 3x3,256 1x1,1024] x 6	[3x3,12] x 12	[Inception] x 2
[3x3,512] x 3		[1x1,512 3x3,512 1x1,2048] x 3		

$$SF = Cat(E, E') \tag{9}$$

In the above equation, *Cat* indicates the concatenation operation.

The detailed architecture of the saliency module with step-by-step operations is illustrated in Fig. 4 (a). The output of every step and its corresponding size have been drawn. The saliency module would output a saliency enhanced feature tensor E and a saliency suppressed feature tensor F'. The architecture of the concatenated saliency module is shown in Fig. 4 (b). The saliency suppressed feature tensor F' is further employed as the input feature of another saliency module and produce another saliency enhanced feature tensor E'. All the saliency enhanced feature tensors are concatenated together to obtain discriminative feature representations of an EL image.

The overall defect detection procedure is summarized in Alg. 1. The basic features of the input image are first extracted via a baseline module. The image features can be first refined by the global pairwise similarity module and then be further refined by the concatenated saliency module to get the final refined features. A global average pooling layer is employed to reduce dimension of the refined features and obtain a feature vector. Then the fully connected layers are applied to predict whether the EL image is defective or not. The proposed method is end-to-end and different modules are trained at the same time.

Algorithm 1 Automatic defect detection for EL images via global pairwise similarity and concatenated saliency

Input: An EL image.
(1) Extract basic features of the input image via a baseline convolutional module, e.g. VGG, shown as the Feature Extraction part of Fig. 2.
(2) First refine image features based on the global pairwise similarity module, shown as the first part of Feature Refining in Fig. 2. The step-by-step computation of this module has be summarized in Fig. 3.
(3) Further refine image features using the concatenated saliency module to get the final refined features, shown as the second part of Feature Refining in Fig. 2. The step-by-step description of this module is illustrated in Fig. 4.
(4) Employ a global average pooling and the fully connected layers as the classifier to predict whether the EL image contains the defect or not, shown as the Defect Classification part of Fig. 2.
Output: The defect detection result of the input EL image.

IV. EXPERIMENTS

Our task is automatic detection of defective photovoltaic module cells with EL images. The proposed global pairwise similarity and concatenated saliency guided neural network is modular and independent of the baseline, which can be easily extended to any convolutional neural network. In order to demonstrate the effectiveness of the proposed network, the performance of the proposed neural network is compared with five popular convolutional neural networks, i.e. VGG16 [18], ResNet56 [16], ResNet50 [16], DenseNet40 [17] and GoogleNet [19]. The configurations of these baseline models are included in Table I. The classifier parts of these baselines are excluded in this table since the baseline models are employed to extract basic features. For GoogleNet, Inception is the basic block, which can be referred in the work [19].

A. Experimental setup

1) Database: The database consists of 2624 EL images which are used for automatic defect recognition [7]. Every EL image is a grayscale image at a resolution of 300x300 pixels. According to the type of the source solar module, these EL images can be divided into two types: monocrystalline solar cell images and polycrystalline ones. The defect probability of each image is during 0 and 1. In our experiments, the probability was binarized with the threshold of 0.5 to get two different labels: 0 (nondefective, i.e. negative samples) and 1 (defective, i.e. positive samples).

75% of the EL images were randomly chosen as training samples and the remaining 25% were chosen for testing. Stratified random sampling was employed to retain the distribution of different types and different labels for the training and testing sets.

2) Parameter setup: Our implementation was based on Pytorch. All our experiments were carried out on a 64-bit Ubuntu OS running with Intel Z390, 32 GB of RAM, and one NVIDIA GeForce GTX 2080Ti with 11GB of RAM.

For VGG16, ResNet56, DenseNet40 and GoogleNet, pretrained weights based on the database of CIFAR-10 were used to finetune different models. For ResNet50, pretrained weights based on the database of ImageNet were employed. Data augmentation was employed to generate additional training samples to avoid potential overfitting. The training samples were randomly flipped along the horizontal and vertical axes. The random scale was limited to the range of $[0.9, 1.1]$ of the original resolution and the random translation was limited to ± 0.1 of the cell dimensions. The training and testing images were resized to have a resolution of 300x300 pixels.

For simplicity, cross-entropy loss was employed as the loss function and SGD optimizer was used with a learning rate of 0.005, weight decay of 5×10^{-4} and a momentum of 0.9. In the training stage, the batch size was 20, which could be larger or smaller depending on the model complexity. The training process was run with 101 epoches in one single step to classify solar module cells. For fairness, each model was trained and tested 7 times and the average performance was adopted to compare different models.

3) Evaluation Metrics: The probability was thresholded with the value 0.5 to get the final classification result. Different models were evaluated using the metrics of Accuracy, F1, Precision and Recall. Precision is the number of true defective samples detected by a model divided by the total number of samples detected. Recall is the number of the true defective

TABLE II
THE PERFORMANCES OF DIFFERENT MODELS, EVALUATED BY AP, AUC, F1, PRECISION, RECALL, BALANCED ACCURACY (B-ACCURACY) AND ACCURACY.

Model	AP	AUC	F1	Precision	Recall	B-Accuracy	Accuracy
VGG16	90.86%	93.58%	80.66%	81.43%	80.18%	85.85%	87.98%
VGG16+proposed SSN	92.09%	94.57%	82.45%	82.19%	82.84%	87.31%	89.01%
ResNet56	88.97%	92.22%	70.72%	58.53%	90.13%	80.17%	76.43%
ResNet56+proposed SSN	92.37%	94.69%	81.14%	75.31%	88.38%	87.43%	87.07%
ResNet50	89.18%	92.38%	77.98%	73.25%	83.61%	84.79%	85.23%
ResNet50+proposed SSN	92.79%	94.86%	84.10%	83.98%	84.52%	88.52%	90.02%
DenseNet40	85.34%	89.63%	69.12%	58.58%	84.94%	78.57%	76.17%
DenseNet40+proposed SSN	92.98%	95.37%	82.76%	77.93%	88.45%	88.46%	88.47%
GoogleNet	89.40%	92.73%	78.06%	74.72%	82.41%	83.57%	86.89%
GoogleNet+proposed SSN	92.62%	95.02%	81.92%	77.29%	87.39%	87.78%	87.92%

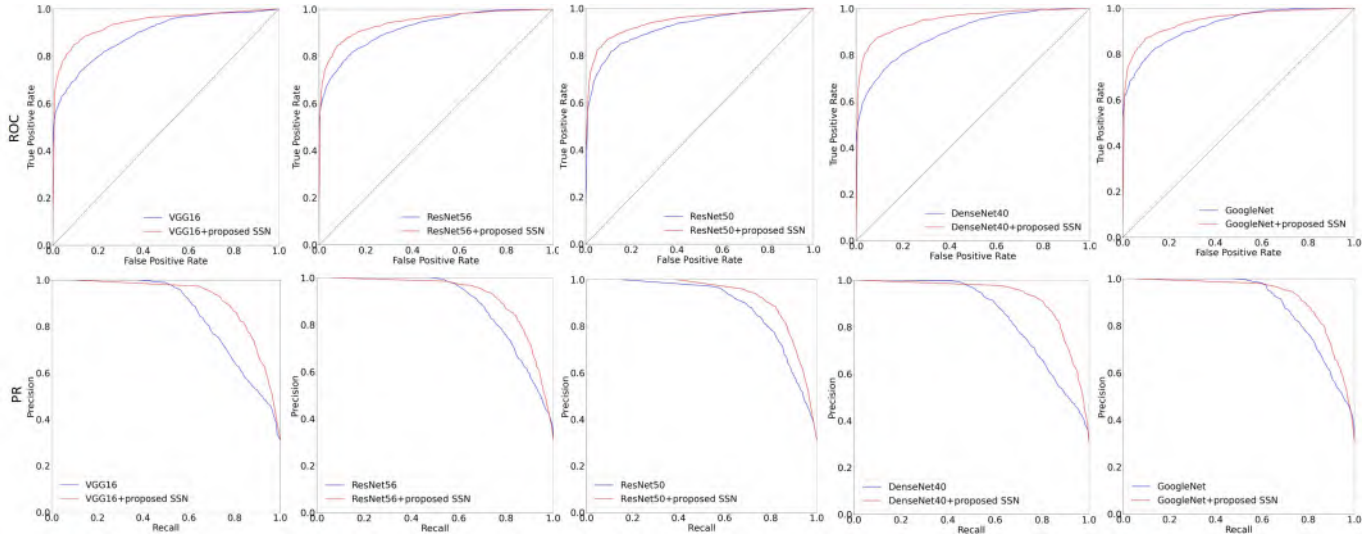

Fig. 5. The ROC curves and PR curves of the proposed SSN network compared with baselines.

samples detected by a model divided by the total number of existing defective samples. Usually, both high precision and high recall are required. As a measure of the overall performance, F1 score is employed, which is calculated as the weighted average of Precision and Recall. Since the database is imbalanced, different models are also compared according to the Balanced Accuracy.

To further understand different models, the probability is thresholded with different values in the range of [0,1] to show the Precision Recall curve (PR curve) and Receiver Operating Characteristic curve (ROC curve). ROC curve has an attractive property: it is insensitive to the imbalance degree of the data [30]. And the Area under the Precision Recall curve (AP) and the Area Under Receiver Operating Characteristic (AUC) are also reported.

B. Performance comparison of different models

To demonstrate the effectiveness of the proposed SSN network, 5 different CNN models are used, i.e. VGG16, ResNet56, ResNet50, DenseNet40 and GoogleNet as the baseline. The scores of AP, AUC, F1, Precision, Recall, Balanced Accuracy and Accuracy of the proposed SSN network and different baseline models are illustrated in Table II. After adopting the proposed SSN network to refine the features extracted by different CNN baseline models, the results prove that the SSN network is good at promoting automatic defect detection performance no matter according to which metric (AP, AUC, F1, Precision, Recall, Balanced Accuracy or Accuracy).

The PR curves and the ROC curves of different models are shown in Fig. 5. The results demonstrate that the proposed SSN network can improve the performance of the baselines with a big gap.

Among the 9 metrics (AP, AUC, F1, Precision, Recall, Balanced Accuracy, Accuracy, the PR curve and the ROC curve) employed in the evaluation, AUC, Balanced Accuracy and the ROC curve are nonsensitive to the distribution of the training samples and thus can provide a fair score for unbalanced data. According to these 3 metrics, the proposed SSN network can better deal with the unbalanced data with performance promotion.

C. performances for mono and poly solar cells

According to the material of solar modules, the solar cells can be divided into two types: monocrystalline silicon and polycrystalline silicon. In EL images, mono solar cell and poly solar cell present different characteristics. As is shown in Fig. 1(a) and Fig. 1(c), the background of mono solar cell EL images is relatively smooth and the background of poly solar cell has a lot of flocculent noises. These different characteristics can produce certain challenge to correctly detect defect. In the proposed SSN network, different types of solar cell images

TABLE III
THE PERFORMANCES OF DIFFERENT MODELS FOR MONO AND POLY SOLAR CELL IMAGES, EVALUATED BY AP, AUC, F1, PRECISION, RECALL, BALANCED ACCURACY (B-ACCURACY) AND ACCURACY.

Type	Model	AP	AUC	F1	Precision	Recall	B-Accuracy	Accuracy
Mono	VGG16	91.39%	93.47%	80.52%	79.50%	82.14%	85.31%	86.30%
	VGG16+proposed SSN	92.54%	94.53%	82.25%	81.25%	83.70%	86.65%	87.58%
	ResNet56	88.94%	91.30%	73.02%	62.57%	88.66%	79.88%	77.13%
	ResNet56+proposed SSN	93.89%	95.35%	81.99%	75.53%	90.22%	87.15%	86.19%
	ResNet50	92.36%	94.33%	79.07%	71.43%	89.44%	84.98%	83.58%
	ResNet50+proposed SSN	94.95%	96.29%	85.91%	84.19%	88.04%	89.60%	90.09%
	DenseNet40	88.03%	90.32%	71.58%	60.72%	88.35%	78.55%	75.48%
	DenseNet40+proposed SSN	94.10%	95.79%	82.47%	75.35%	91.46%	87.73%	86.57%
	GoogleNet	88.41%	91.14%	75.81%	71.11%	82.01%	81.22%	84.38%
	GoogleNet+proposed SSN	93.49%	95.27%	80.81%	73.50%	90.22%	86.38%	85.18%
Poly	VGG16	90.74%	93.93%	80.83%	83.53%	78.57%	86.02%	89.15%
	VGG16+proposed SSN	91.99%	94.82%	82.62%	83.20%	82.14%	87.66%	89.97%
	ResNet56	89.44%	93.03%	69.06%	55.81%	91.32%	80.49%	75.94%
	ResNet56+proposed SSN	91.32%	94.31%	80.52%	75.56%	86.86%	87.44%	87.68%
	ResNet50	87.03%	91.71%	77.05%	75.47%	78.82%	84.15%	86.38%
	ResNet50+proposed SSN	91.07%	93.88%	82.59%	83.90%	81.63%	87.51%	89.97%
	DenseNet40	83.32%	89.23%	67.15%	57.13%	82.14%	78.27%	76.65%
	DenseNet40+proposed SSN	92.23%	95.16%	83.06%	80.48%	85.97%	88.66%	89.79%
	GoogleNet	90.28%	93.75%	79.88%	77.68%	82.86%	85.40%	88.64%
	GoogleNet+proposed SSN	92.25%	95.00%	82.97%	81.28%	85.08%	88.42%	89.82%

TABLE IV
ABLATION STUDY OF THE PROPOSED NETWORK. VGG16 IS THE BASELINE. GS INDICATES THE GLOBAL PAIRWISE SIMILARITY MODULE AND CS IS THE CONCATENATED SALIENCY MODULE. THE PROPOSED NETWORK IS VGG16+GS+CS

Model	AP	AUC	F1	Precision	Recall	B-Accuracy	Accuracy
VGG16	90.86%	93.58%	80.66%	81.43%	80.18%	85.85%	87.98%
VGG16+GS	91.56%	94.12%	81.35%	80.86%	82.00%	86.55%	88.26%
VGG16+CS	90.99%	93.65%	81.32%	82.38%	80.39%	86.27%	88.48%
VGG16+GS+CS	92.09%	94.57%	82.45%	82.19%	82.84%	87.31%	89.01%

TABLE V
THE PERFORMANCES OF THE VARIANT MODELS BASED ON THE BASELINE OF VGG16 WHEN COMPUTING THE SALIENCY SENSITIVE WEIGHTS ALONG THE HEIGHT OF AN IMAGE (SALH), ALONG THE WIDTH OF AN IMAGE (SALW) AND ALONG BOTH THE HEIGHT AND WIDTH (SALHW)

Metric	salH	salW	salHW
Accuracy	89.01%	87.30%	87.49%
B-Accuracy	87.31%	86.96%	87.08%

are not separately trained and it is interesting to see whether the proposed network can well generalize to different types of solar cells.

The performances of different models are quantitatively calculated for mono solar cell images and poly solar cell images, which can be seen in Table III. No matter based on which baseline, all the metrics (AP, AUC, F1, Precision, Recall, Balanced Accuracy and Accuracy) indicate that the proposed SSN network can well refine both the features of mono solar cell images and the features of poly solar cell images and thus get a performance promotion for both types.

D. Ablation study

The proposed SSN network is consisted of two modules: a Global pairwise Similarity module and a Concatenated Saliency module, which are abbreviated as GS and CS separately for simplicity. An ablation study is performed to evaluate each component of the proposed SSN network. Based on the baseline of VGG16, the global pairwise similarity module is added to get a temp model, which is indicated as VGG16+GS. The temp model which only combines the concatenated saliency module is indicated as VGG16+CS. And the final model, which combines both the global pairwise similarity module and the concatenated saliency module, is indicated as VGG16+GS+CS (i.e. VGG16+proposed SSN), for clarity. The results of different models are shown in Table IV. According to multiple different metrics (AP, AUC, F1, Precision, Recall, Balanced Accuracy and Accuracy), the final model, i.e. VGG16+GS+CS, has the best overall performance. The temp model with a global pairwise similarity module (VGG16+GS) and the temp model with a concatenated saliency module (VGG16+CS) have the median performances. And the baseline, i.e. VGG16 has the worst performance. The experimental results also show that VGG16+GS helps to boost the metric of Recall, which means that the global pairwise similarity module can help to detect out the existing defective images. The reason might be that the global pairwise similarity module models the long dependencies and can extract the features of long and thin objects which may disappear or become blurry in the deep layers of the baseline modules. VGG16+CS helps to improve the performance of Precision, which indicates that the concatenated saliency module can help to improve the accuracy of the samples detected as defective. The reason might be that the concatenated saliency module can effectively suppress the background and decouple the true defective part and the noises in the image background, such as the horizontal gratings and irregular floccules. Thus, these noises would produce fewer disturbances in the defect detection tasks. The above comparisons and observations all demonstrate the effectiveness of each component for the proposed SSN network.

As is shown in Fig. 4(a), the saliency sensitive weights are computed along the height of an image. It is curious to see how the proposed network works when computing the saliency sensitive weights along the width of an image and along both the height and width. These different ways are abbreviated as salH (along the height), salW (along the width), and salHW (along both the height and width) separately for simplicity. The results are shown in Table V.

From the table, it can be observed that computing the

TABLE VI
THE PERFORMANCES OF THE PROPOSED MODEL WITH DOT PRODUCT SIMILARITY, COSINE SIMILARITY AND PEARSON CORRELATION COEFFICIENT BASED ON THE BASELINE OF DENSENET40

Metric	dot product	cosine	Pearson
Accuracy	88.47%	89.41%	88.34%
B-Accuracy	88.46%	88.56%	87.94%

saliency sensitive weights along the height of an image has the best performance and computing the saliency sensitive weights along the width of an image has the worst performance. The reason might be that there are horizontal gratings in EL images and computing the saliency sensitive weights along the height of an image would consider each grating as a whole, which makes sense.

In all the above experiments, the dot product similarity is employed to measure the similarity between the image pixel and the representative patch in the global similarity module, introduced in 1. In this section, different other methods, i.e. cosine similarity and Pearson correlation coefficient are further tested to measure the similarity. Table VI shows that cosine similarity has the best performance while Pearson correlation coefficient has the worst performance, which indicate that introducing or designing a proper method to measure the similarity may further improve the performance of the proposed method.

E. Impact of the number of training samples

Because the data is unbalanced, training images are randomly sampled with weights. The weight for each class (defective class and nondefective class) is the reciprocal of the image number in the corresponding class of the training set. And the number of training images randomly sampled is set as 3000. This section aims to see how the number of training samples influences the result of the proposed network. The performances of the proposed network with 1500 training samples, 3000 training samples and 6000 training samples based on the baseline of DenseNet40 are shown in Table VII. It can be observed that larger number of training samples would make the proposed network have a better performance. In the above all experiments, the number of training images is set as 3000 to save training time. And the performance of the proposed SSN network can be further improved by using more training samples.

F. Train loss vs validation loss

The train and validation loss of the proposed SSN network with DenseNet40 as the backbone are plotted in Fig. 6. The diagram shows that the validation loss became smallest at about epoch 60 and the train loss became smallest at approximately epoch 96. These observations also show that a small number of samples is easy to overfit and a large number of training samples may help to further improve the performance.

G. Discussion

According to the above observations of the experiments, there are multiple ways to further improve the defect detection

TABLE VII
THE PERFORMANCES OF THE PROPOSED MODEL WITH 1500 SAMPLES, 3000 SAMPLES AND 6000 SAMPLES BASED ON THE BASELINE OF DENSENET40

Metric	1500	3000	6000
Accuracy	86.92%	88.47%	89.67%
B-Accuracy	87.87%	88.46%	88.80%

Fig. 6. The train and validation loss of the proposed network with DenseNet40 as backbone.

performance of the proposed network. Firstly, the performance can be further improved with a larger number of training images. Secondly, the proposed network can be further improved by introducing a better method to compute the similarity between pixels and patches in the global similarity module. The third way to improve the performance is to separately train two different types of solar cell images: monocrystalline silicon and polycrystalline silicon. In the proposed SSN network, different types of solar cell images are not separately trained and the experiment results show that the proposed model can improve the detection performance for both mono solar cell images and poly solar cell images. However, if separately training for mono and poly images, the different characteristics of these two types of images are clearly separated in advance and the performance of defect detection might be further enhanced.

Despite the good performance based on the above experiments, there are also shortcomings of the proposed SSN network. Based on these shortcomings, there are more ways to improve the performance. The main shortcoming of the proposed method is that if the defect is small and very similar to the surrounding background, it is hard for the proposed method to correctly detect. Actually, small and camouflaged object detection task is very hard even for natural scene images. There have been a lot of researchers working on this task. Introducing useful methods to extract proper features for small and camouflaged objects would help to improve the defect detection performance. Another shortcoming of the proposed method is that the proposed method does not deal with the uncertainties of the labels created by experts. The defect detection task can be hard even for an expert, thus there usually are noises in the data labels. If the uncertainties of the labels can be further considered, the performance could be further improved.

V. CONCLUSION

This paper focuses on automatic defect detection for photovoltaic module cells in EL images. The main contribution of this paper is that a global pairwise similarity and concatenated saliency guided neural network is proposed to refine features extracted by the convolutional neural network to automatically

detect defect. The proposed network explicitly considers the visual characteristics of the defects and the noises in solar cells. Since the defects in solar cells are usually long and thin, a global pairwise similarity module is exploited to well model long-range dependencies in images. Another observation of the solar module cell is that there are a lot of noises in the EL images, such as the horizontal gratings and irregular floccules. Based on this observation, a concatenated saliency module is exploited to suppress the background and decouple different salient features to better represent a solar cell image. Extensive experiments based on 5 different baselines (VGG16, ResNet56, ResNet50, DenseNet40 and GoogleNet) demonstrate that the proposed network can well refine the features and improve the defect detection performance for both mono and poly solar cell images. Also, ablation studies prove that both the global pairwise similarity module and the concatenated saliency module are effective.

Despite the good performance of the proposed SSN network, future research is recommended on the following topics.

1) More training samples: There are multiple ways to introduce more training samples, such as collecting more defective images, employing useful sampling techniques or generating samples based on Generative Adversarial Networks (GAN).

2) Small and camouflaged defects detection. The performance might be further improved by constructing an appropriate way to extract multi-scale features and model the information of the context.

3) Modeling uncertainties. This research could be realized by modeling the uncertainty of the image sample and introducing a way to exclude the interference of the uncertain samples.

REFERENCES

[1] K. Burrows and V. Fthenakis, "Glass needs for a growing photovoltaics industry," *Solar Energy Materials and Solar Cells*, vol. 132, pp. 455–459, 2015.

[2] S. Osawa, T. Nakano, S. Matsumoto, N. Katayama, Y. Saka, and H. Sato, "Fault diagnosis of photovoltaic modules using ac impedance spectroscopy," in *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*. IEEE, 2016, pp. 210–215.

[3] C. Li, Y. Yang, S. Spataru, K. Zhang, and H. Wei, "A robust parametrization method of photovoltaic modules for enhancing one-diode model accuracy under varying operating conditions," *Renewable Energy*, vol. 168, pp. 764–778, 2021.

[4] K. Bedrich, M. Bokalič, M. Bliss, M. Topič, T. R. Betts, and R. Gottschalg, "Electroluminescence imaging of pv devices: advanced vignetting calibration," *IEEE Journal of Photovoltaics*, vol. 8, no. 5, pp. 1297–1304, 2018.

[5] J. Coello, "Introducing electroluminescence technique in the quality control of large pv plants," in *26th European Photovoltaic Solar Energy Conference and Exhibition*, 2011, pp. 3469–3472.

[6] C. Lupangu and R. Bansal, "A review of technical issues on the development of solar photovoltaic systems," *Renewable and Sustainable Energy Reviews*, vol. 73, pp. 950–965, 2017.

[7] 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*, vol. 185, pp. 455–468, 2019.

[8] A. M. Karimi, J. S. Fada, M. A. Hossain, S. Yang, T. J. Peshek, J. L. Braid, and R. H. French, "Automated pipeline for photovoltaic module electroluminescence image processing and degradation feature classification," *IEEE Journal of Photovoltaics*, vol. 9, no. 5, pp. 1324–1335, 2019.

[9] 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*, vol. 99, pp. 250–262, 2012.

[10] M. Dhimish and V. Holmes, "Solar cells micro crack detection technique using state-of-the-art electroluminescence imaging," *Journal of Science: Advanced Materials and Devices*, vol. 4, no. 4, pp. 499–508, 2019.

[11] Y. Wang, L. Li, Y. Sun, J. Xu, Y. Jia, J. Hong, X. Hu, G. Weng, X. Luo, S. Chen *et al.*, "Adaptive automatic solar cell defect detection and classification based on absolute electroluminescence imaging," *Energy*, vol. 229, p. 120606, 2021.

[12] M. Dhimish, V. Holmes, and P. Mather, "Novel photovoltaic micro crack detection technique," *IEEE Transactions on Device and Materials Reliability*, vol. 19, no. 2, pp. 304–312, 2019.

[13] D.-M. Tsai, S.-C. Wu, and W.-Y. Chiu, "Defect detection in solar modules using ica basis images," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 122–131, 2012.

[14] B. Su, H. Chen, Y. Zhu, W. Liu, and K. Liu, "Classification of manufacturing defects in multicrystalline solar cells with novel feature descriptor," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 12, pp. 4675–4688, 2019.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, 2012, pp. 1097–1105.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700–4708.

[18] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv:1409.1556*, 2014.

[19] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 1–9.

[20] C. Ge, Z. Liu, L. Fang, H. Ling, A. Zhang, and C. Yin, "A hybrid fuzzy convolutional neural network based mechanism for photovoltaic cell defect detection with electroluminescence images," *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 7, pp. 1653–1664, 2021.

[21] C. Shou, L. Hong, W. Ding, Q. Shen, W. Zhou, Y. Jiang, and C. Zhao, "Defect detection with generative adversarial networks for electroluminescence images of solar cells," in *2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*. IEEE, 2020, pp. 312–317.

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

[23] M. W. Akram, G. Li, Y. Jin, X. Chen, C. Zhu, X. Zhao, A. Khaliq, M. Faheem, and A. Ahmad, "Cnn based automatic detection of photovoltaic cell defects in electroluminescence images," *Energy*, vol. 189, p. 116319, 2019.

[24] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in Neural Information Processing Systems*, 2015, pp. 91–99.

[25] B. Su, H. Chen, P. Chen, G. Bian, K. Liu, and W. Liu, "Deep learning-based solar-cell manufacturing defect detection with complementary attention network," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 6, pp. 4084–4095, 2021.

[26] C. Fang, H. Tian, D. Zhang, Q. Zhang, J. Han, and J. Han, "Densely nested top-down flows for salient object detection," *arXiv preprint arXiv:2102.09133*, 2021.

[27] D. Zhang, J. Han, Y. Zhang, and D. Xu, "Synthesizing supervision for learning deep saliency network without human annotation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 42, no. 7, pp. 1755–1769, 2019.

[28] X. Chen, C. Fu, Y. Zhao, F. Zheng, J. Song, R. Ji, and Y. Yang, "Salience-guided cascaded suppression network for person re-identification," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 3300–3310.

[29] J. Zhang, K. A. Ehinger, H. Wei, K. Zhang, and J. Yang, "A novel graph-based optimization framework for salient object detection," *Pattern Recognition*, vol. 64, pp. 39–50, 2017.

[30] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

Jinxia Zhang received the B.S. degree in the Department of Computer Science and Engineering, Nanjing University of Science and Technology, China in 2009 and the PhD degree in the Department of Computer Science and Engineering, Nanjing University of Science and Technology, China in 2015. She was a visiting scholar in Visual Attention Lab at Brigham and Womens Hospital and Harvard Medical School from 2012 to 2014. She is currently an associate professor in the School of Automation, Southeast University. Her research interests include saliency detection, knowledge transfer, computer vision and machine learning.

Yu Shen received the B.S. degree in the School of Internet of Things, Jiangnan University, China in 2013 and the Master degree in the School of Automation, Southeast University, China in 2016. She is currently the Ph.D. candidate in the School of Automation, Southeast University. Her research interests include solar energy, photovoltaics, computer vision and machine learning.

Jiacheng Jiang received the B.S. degree in Automation from Jiangnan University, in 2020, and is currently working toward the M.S. degree in Electronic and Information Engineering at Southeast University. His current research interests include image classification and image segmentation.His recent work has focused on image classification and defect inspection.

Shixiong Fang received the B.S. degree in the School of Electrical and Electronic Engineering, Hubei University of Technology, China in 2009, and the M.S. and Ph.D. degrees in control theory and control engineering from Southeast University, China in 2002 and 2009. He was a visiting scholar in European Organization for Nuclear Research from 2004 to 2006. He is currently a lecturer in the School of Automation, Southeast University. His research interests include computer vision and cyber physical systems.

Liping Chen received Master of Optical Engineering, MBA of Tongji University, Senior Engineer, Vice Principal of Jiangsu (Suntech) Institute for Photovoltaic Technology, and off-campus cooperative advisor of graduate students with masters degrees in School of Science, Jiangnan University. She joined Suntech (Wuxi) in August 2005 and has been engaged in research and development and technical management of high-efficiency solar cells for more than 16 years. During her work, she participated in eight national-level, provincial-level, and international cooperative scientific research projects as the core technical backbone, applied for 24 patents, and applied for 11 patents as the first inventor, which had four invention patents and two utility model patents authorized by the China National Intellectual Property Administration. She was awarded the title of "Young and Middle-aged Experts with Outstanding Contributions in Wuxi City" in 2020, and won the first prize of Jiangsu Photovoltaic Science and Technology Award in 2020 and the second prize in the Science and Technology Progress Award of Chinese Renewable Energy Society in 2020.

Tingting Yan received the M.S. degree in Optical Engineering, Senior Engineer, Patent Attorney, and Off-Campus Academic Advisor for Postgraduates of Jiangnan University. She joined Wuxi Suntech in 2009 and engaged in the research and development of high-efficiency crystalline silicon solar cells. During this period, participated in 4 national, provincial and ministerial projects, and won the Wuxi Patent Award and the Jiangsu Science and Technology Award and other honors; currently engaged in intellectual property management, industry standardization, and industry-university-research project cooperation in the Institute.

Zuoyong Li received the B.S. and M.S. degrees in Computer Science and Technology from Fuzhou University, Fuzhou, China, in 2002 and 2006, respectively. He received the Ph.D. degree from the School of Computer Science and Technology at Nanjing University of Science and Technology, Nanjing (NUST), China, in 2010. He is currently a professor in College of Computer and Control Engineering at Minjiang University, Fuzhou, China. He has published over 70 papers in international/national journals. His current research interest is image processing, pattern recognition and machine learning.

Kanjian Zhang received the B.S. degree in mathematics from Nankai University, China in 1994, and the M.S. and Ph.D. degrees in control theory and control engineering from Southeast University, China in 1997 and 2000. He is currently a professor in the School of Automation, Southeast University. His research is in nonlinear control theory and its applications, with particular interest in robust output feedback design and optimization control.

Haikun Wei received the B.S. degree in the department of Automation, North China University of Technology, China in 1994, and the M.S. and Ph.D. degrees in the Research Institute of Automation, Southeast University, China in 1997 and 2000. He was a visiting scholar in RIKEN Brain Science Institute, Japan from 2005 to 2007. He is currently a professor in the School of Automation, Southeast University. His research interest is real and artificial in neural networks and industry automation.

Weili Guo received the B.S. degree in School of Science, Shandong Jianzhu University, China in 2007, the M.S. degree in School of Science, Nanjing Agricultural University, China in 2010, and the Ph.D. degree in School of Automation, Southeast University, China in 2014. From 2016 to 2017, he was a postdoctoral fellow in School of Computer Science and Engineering, Nanyang Technological University, Singapore. He is currently working as an Associate Professor with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His research interests include neural networks, continual learning and artificial intelligence.