Classification of Defects in Photovoltaic Cells from Electroluminescence Images using Advanced Deep Learning Techniques

Smita D Khandagale
Department of Electronics Engineering
Datta Meghe College of Engineering, Airoli, Navi-Mumbai
University of Mumbai, MS, India
sdkhandagale@vpmthane.org

Sanjay M. Patil
Department of AIDS Engineering
Datta Meghe College of Engineering, Airoli, Navi-Mumbai
University of Mumbai, MS, India
sanjay.patil@dmce.ac.in

2024 2nd DMIHER International Conference on Artificial Intelligence in Healthcare, Education and Industry (IDICAIEI) | 979-8-3315-2871-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/IDICAIEI61867.2024.10842697

***Abstract*—For a long time, keeping big solar power plants running has been a real challenge. This work talks about an exciting new way to detect problems in solar panels using special pictures called electroluminescence images (EL). Deep learning algorithms is used to help with this task! Our research introduces a fresh method that can automatically spot flaws by using a pre-trained deep learning Vision Transformer (ViT) model. The model is trained on ELPV Dataset with EL image, which is available for everyone. The model achieved an impressive accuracy of 94.53% on the test set! Our model was thoroughly tested against popular machine learning models like ResNet50, DenseNet, ResNet101, VGG-16, VGG-19, Inception ResNet, ResNet-152, MobileNet, and Xception. The results showed that proposed model did better than all of these earlier models. The model is compared with other top methods for finding problems in EL images. Proposed method has proved to be better than the existing models. It really showed it can accurately find flaws in electroluminescence images.**

Keywords—Deep learning, Photovoltaic cell images, Solar Cell, Vision Transformer, Electroluminescence

I. INTRODUCTION

Renewable energies are sources of energy that are both clean and free from carbon emissions and greenhouse gases. According to the International Energy Agency (IEA), the demand for electricity is projected to rise by 70% by 2040, mostly due to the growth in India, China, South-East Asia, and the Middle East [1]. Solar energy has the potential to greatly diminish reliance on fossil fuels. The government and commercial organisations are motivated to develop solar energy setups due to the quantity of sunshine, simplicity of operation, cost-effectiveness, and safety. Most countries plan to use solar energy as a source of energy after the recent energy conference substitute for non-renewable energy sources such as coal, with the aim of diminishing their dependence on them. Photovoltaic cells, sometimes referred to as Photovoltaic (PV) cells, produce electrical energy by converting photons into electrons via the absorption of light energy. A lot of the time, PV cells are put into two groups based on their crystal structure: mono-crystalline and poly-crystalline. Monocrystalline cells, which are made up of just one crystal, make more energy than other types. Conversely, polycrystalline PVs, composed of numerous shards of crystalline material, have lower efficiency and have faster fault development [2]. Periodic and automatic categorization and grading of photovoltaic cells are essential for accurately quantifying and predicting the energy and derived from the unique characteristics of Monocrystalline (MC), Polycrystalline (PC), and Photovoltaics (PVs). Moreover, the power output is greatly influenced by the quality used for photovoltaic materials and the presence of flaws.

Nevertheless, PV cells are prone to several sorts of defects and fractures that might occur during production, installation, transit, or operation. These problems may result in decreased efficiency, causing challenges in load planning and perhaps leading to power shortages that might jeopardise the industrial sector [3]. These challenges need precise, immediate, non-destructive identification of flaws and fractures in the PV cells to ensure the proper functioning of solar power producing facilities.

An analysis of the flaws in the PV cells may be conducted using measurements of electrical voltage, current, power, as well as thermal and visual pictures [4]. Nevertheless, the electrical characteristics often fail to identify minuscule and microscopic fractures or other imperfections, since these tiny cracks do not cause any discernible alteration [5]. In addition, electrical measurements are unable to accurately identify the flaws and can only be used for individual PV cells, making it impractical in real-world scenarios. However, it should be noted that thermal imaging-based approaches may not always be reliable as high temperatures may not always indicate the presence of fractures [6,7]. The EL imaging technique uses a charged coupled device to capture pictures within the electromagnetic wavelength range of 950-1200nm, while operating in the forward bias state. Typically, EL imaging enhances the visibility of faulty areas by making them look darker, which would be difficult to detect without the use of visual aids. Regular physical inspections by the maintenance personnel are essential for making crucial operational decisions at these expansive solar power producing installations. The PV cells are either kept or replaced dependent on their physical state. Nevertheless, the distant positioning of these units necessitates tedious and expensive occasional human inspections [8-10].

This paper talks about deep learning model called the ViT. It can find defects in solar cells really well by looking at electroluminescence (EL) images. The Transformer is based on the Transformer model. Now, the Transformer has been getting a lot of attention lately. Results in tasks like machine translation and things related to Natural Language Processing (NLP). This model works using an encoder-decoder framework. One neat thing is that it can handle sequences of input all at once. That means it doesn't need to rely on those slower recurrent neural networks anymore.

Objective of the work:

The work includes a pre-trained Vision Transformer (ViT), a sophisticated machine learning model specifically created to automatically categorise EL photos from the ELPV dataset.

Ultimately, the model was subjected to cross-validation and testing using the photographs.

The pre-trained ViT model is assessed based on many performance indicators, including precision, F1-score, accuracy, and sensitivity. And scores that have been assigned a weight. In addition, Comparison between suggested model and nine other pre-trained deep learning models, as well as the currently leading models in the field.

The paper is constructed as follows: Section II provides literature Survey, a detailed explanation of the technological framework of the Pre-trained ViT model is provided in Section III. The experimental findings are analysed and deliberated upon in part IV. The article is concluded in Section V.

II. LITERATURE REVIEW

Presently, there is a significant surge in the popularity of deep learning models that have been trained on large datasets. Specifically, Convolutional Neural Networks (CNN) are extensively used in several fields like object recognition, picture classification, and semantic segmentation. Currently, the process of identifying surface defects on solar cells is increasingly being achieved by deep learning, which has important implications for the advancement of intelligent manufacturing. Nevertheless, deep learning algorithms are primarily intended for natural scene photos, and using them directly to identify surface defects in solar cell EL images poses many difficulties. Therefore, researchers must use task-specific approaches, such as data processing, feature engineering, and creative neural network topologies, to effectively tackle the intricacy and distinctiveness of solar cell defect detection. Defect detection techniques that use deep learning may be classified into three categories based on the sorts of industrial inspection tasks: segmentation networks, detection networks, and classification networks.

Deitsch et al. [13] introduced two deep learning approaches, including Support Vector Machines (SVM) and CNN, to automatically identify errors in PV cells. The experimental findings demonstrated that the CNN classifier had a remarkable level of accuracy in detecting defects. Pierdicca [14] used transfer learning using the VGG-16 network to categorise solar cell photos obtained from remote sensing. Nevertheless, the CNN network achieves an accuracy of around 70% owing to the reduced picture resolution in the self-constructed solar module electroluminescence image collection. Tang et al. [15] developed a CNN model to automatically classify EL picture faults. The CNN extracts deep characteristics from the input and feeds them into fully connected layers for the purpose of classifying pictures into four fault groups. Nevertheless, the model just ascertains the presence of flaws and is unable to pinpoint the exact locations or kinds of problems. Sridhar et al. [16] performed data augmentation on PV photos obtained by unmanned aerial vehicles in order to increase the size of the dataset. A CNN model was used to categorise data into five different fault classes and a category representing samples without any defects. Their model attained a really high degree of precision. Korkmaz et al. [17] used an existing architecture to create a new multi-scale model for identifying different flaws in solar panels. This technique demonstrated exceptional resilience and achieved outstanding classification accuracy.

Su et al. [18] introduced a new object detector for identifying defects in solar cells. They included a specially constructed bidirectional feature pyramid into the model. This facilitated the efficient identification and detection of concealed fractures, grid disruptions, and flaws characterized by dark spots. Nevertheless, the method still needs human tweaking of the feature balancing factor. Zhang et al. [19] developed a multi feature area proposal fusion network structure to identify concealed fractures and grid break flaws in polycrystalline solar panels. This network uses CNN to extract area suggestions from several feature layers. However, the model incurs significant computing expenses and results in lengthy detection durations. Xu et al. [20] used a novel spatial pyramid pooling technique and channel attention into the YOLOv5 model to detect fractures and fragment flaws in EL pictures. Chen et al. [21] developed an innovative defect object detector that integrates a dual-channel feature pyramid into YOLOv5, enhancing the model's capacity to identify minor target flaws. Nevertheless, the model has a limited ability to identify a smaller range of flaws in solar cells. Balcıoğlu et al. [22] developed a novel Deep CNN to create a visual defect detection model. During the first phase, it identifies solar cell samples that exhibit flaws and categorises them according to the extent of their damage. During the second step, the chosen samples are categorised, hence enhancing the detection accuracy for minor errors in intricate backgrounds. Nevertheless, the resolution of photographs in their collection is quite poor owing to cost constraints.

III. METHODOLOGY

A. *Defect Analysis Dataset for Solar Photovoltaic Cell*

A set of freely available data about solar cells was used to test the pre-trained ViT model. High-resolution 224×224 Electroluminescence (EL) pictures from 44 PV solar panels are in the collection. There are 2624 EL shots in all, which have been divided into two groups: monocrystalline PV modules and polycrystalline PV modules. The photographs were taken in a dimly lit area to ensure consistent lighting, since photovoltaic modules only release light during the capturing process. The solar cell pictures that depict the subject are shown in Figure 1. The retrieved cell pictures were also shown in a random manner to an expert, with the primary emphasis being on flaws where the power loss exceeded 3%. The evaluations of operational and malfunctioning cells, as well as those that did not meet the required criteria, were categorised as faulty. Control-rate's uncertainty was given a lower weight, with 33% allocated to non-confident evaluation of functioning cells and 67% assigned to non-confident assessment of faulty cells.

B. *Model Proposal: ViTs with pre-training*

The ViT has been suggested as a way to make the regular Transformer model more useful for the job of classifying images. The main goal is to make it work for all kinds of conversation, not just written words, without using any special data structures. The decoder section of the Transformer is used by the ViT method to do classification work. This is done by connecting a set of picture patches to the semantic explanation that goes with them. The attention mechanism of the ViT enables selective concentration on various regions of an image and integration of data from the whole visual content. This is different from traditional CNN designs that usually use filters with a limited scope. Figure 2 depicts the all-inclusive architecture of the model from start to finish. At first, a training set picture X is split up into pieces that don't cross.

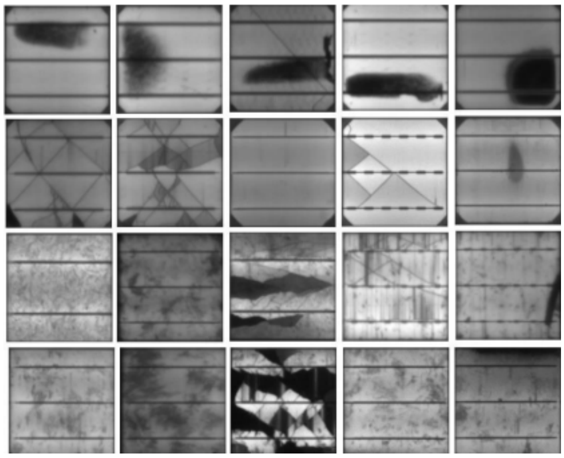

Fig. 1: Displayed are sample pictures extracted from the

ELPV dataset. The photos in the first row are monochromatic, while the images in the second row are polychromatic

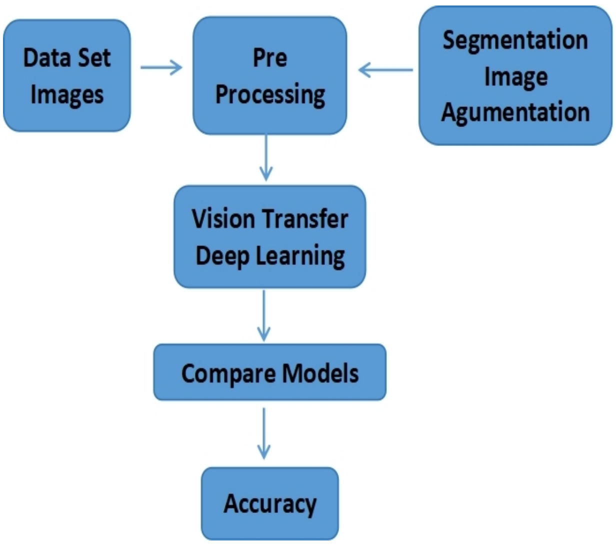

Fig. 2: Proposed Model for the Defect detection

C. *Selecting the ViT-Base model*

The experimental findings, obtained by using ViTs of varying sizes, indicate that more precision may be attained by utilising models with increased depth. Another benefit of choosing a smaller patch size is that it increases the sequence length,. This makes the model more accurate overall. The level of skill being referred to shows a direct relationship with the level of model complexity. Unlike models that depend on CNN, which are limited to recognising local information at lower levels, higher layers in the network are required to detect global information. The aforementioned attribute of the ViT is very crucial in discerning pertinent aspects for categorization purposes. ViT-Base with 80 million factors, on the other hand, is the model chosen.

D. *Pretrained model*

Oxford University's Visual Geometry Group (VGG) developed the VGG-16 model, an architecture for a CNN. The depth of the model is defined by its sixteen layers, of which thirteen are convolutional and three are completely linked. VGG-16's simplicity and efficacy have earned it a reputation for good performance on a variety of computer vision tasks, such as object identification and picture categorization. An architecture with gradually increasing depth is included in the model, with a stack of convolutional layers followed by max-pooling layers. The model is able to learn complex hierarchical representations of visual characteristics because to its architecture, which results in predictions that are both resilient and accurate. The adaptability and outstanding performance of VGG-16 continue to make it a popular option for many deep learning applications, even if it is simpler than more modern designs.

ResNet: In conventional convolutional networks, the l^{kh} and $(l+l)^{kh}$ layers are connected to each other by a single link in each layer. Consequently, this results from a straightforward layer transition of

$$X_l = H_l(X\_\{l-1\}) \text{ ----1}$$

The procedure $H_l(\cdot\ )$ is often a pooling or convolution (assume it's subscript). On the other hand, ResNets use skip connections to generate an identity mapping without using the non-linear transformation. Here it is shown.

$$X_l = H_l(X_{l-1}) + X_{l-1} \text{ ---2}$$

By using these skip connections, gradients are able to flow directly between layers. Although this might alleviate the issue of the disappearing gradient, it could also lead to other complications.

DenseNet: If we had an L-layer network, a ResNet would have 2L connections, while a conventional convolutional network would have L. Currently, the DenseNet has L(L+1)/2 connections since there are direct connections between each layer and all subsequent levels. In contrast to ResNets, DenseNets effectively conserve information by concatenating the identity mappings.

$$X_l = H_l(X_0, X_1, \ldots X_{l-1}) \text{ ---3}$$

To rephrase, the feature maps from all the layers before it are combined and sent into layer l, which is responsible for applying the non-linear transformation $H_l(\cdot)$. The DenseNet provides enhanced performance as an enhancement to the classic convolutional network. DenseNets use dense connection to ensure that feature maps are sent down from all layers above it. In addition to enhancing gradient flow and general information flow in the network, this also makes feature reuse easier. Parameter efficiency is further improved by adding bottleneck and transition layers. Not only are short skip connections abundant, but the vanishing gradient issue is also much reduced.

IV. EXPERIMENT RESULTS & DISCUSSION

A. *Methodology and Assessment of Performance*

The ViT-Base model, which was suggested, was developed using the TensorFlow and Keras packages. The ViT model is pre-trained using the ADAM optimizer, with a learning rate of 10^-3 and a batch size of 64 for a total of 100 epochs. Furthermore, every layer in the model may be modified. Three sets were made from the dataset: a set for training with 70% of the pictures, a test set alongside 20% of the pictures, and a validation set about 10% of the pictures. During the training, testing, and approval steps, the picture was always shrunk to 224 by 224 pixels. The typical metrics used for evaluating performance in EL image assessment include accuracy, weighted precision, macro precision, recall, and F1 score. These metrics are used to ensure consistency and analyse performance in multi-class scenarios.

B. Results of image classification using EL images

The pre-trained ViT model was able to find flaws in EL pictures from the ELPV dataset with a high level of accuracy (Table I).

TABLE I. PERCENTAGE OF ERROR METRIC (PEM) SCORES FOR THE PROPOSED PRE-TRAINED VIT (VIT) TABLE TYPE STYLES

Accuracy	***Mean***	***F1-Score***	***Recall***	***Precision***
94.53	Macro	95.2	94.01	94.89
	Weighted	94.66	93.69	94.23

C. Comparative assessment

In the field of Deep Learning, pre-trained networks are a popular way to sort images into groups. They use cutting edge methods to compare the suggested pre trained ViT model to other models.

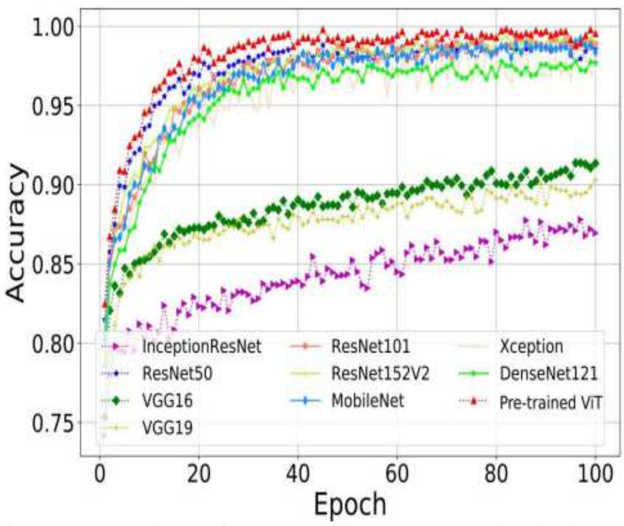

Fig. 3: Comparison of accuracy between pre-trained ViT and pre-trained CNN-based models throughout epochs.

D. Comparison of Pre-trained Models

A large library with more than a million pictures and 1000 groups is usually used to train networks that have already been set up. So, these pre-trained models have a lot of useful qualities because they were trained on big, widely-used picture datasets that contain a lot of different types. The deeper layers of these networks have been improved to adjust and include new characteristics that are important for achieving to classification goal.

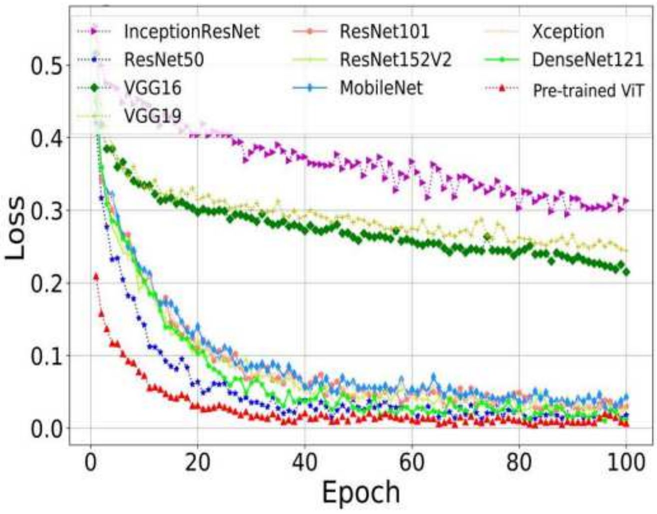

Fig. 4 : Comparison of loss across epochs between pre-trained ViT and pre-trained models based on CNN.

Figure 3 and 4 illustrate the fluctuations in model accuracy and loss across several pre-trained models throughout the training phase on the EL dataset. The pre trained ViT model demonstrates enhanced performance and higher adaptability compared to other models.

TABLE II. COMPARING THE RESULTS WITH NEURAL NETWORKS.

Model	**Accuracy**
InceptionResNet	86.31
ResNet50	92.36
VGG-16	91.54
VGG-19	90.36
DenseNet	89.58
Xception	93.48
MobileNet	93.98
ResNet152	92.78
ResNet101	92.76
Proposed Model	94.53

The EL dataset's classification job is summed up in Table II, which also shows the accuracy scores. ViT that has already been trained does better at Entity Linking than other pre-trained networks, as shown in the last row of Table II.

V. CONCLUSION

This work shows a special pre-trained ViT model that can correctly sort EL images from the ELPV dataset. With an average test success of 94.53%, The ViT model with pre-trained weights and bias values has generated a good accuracy of 94.53% compared with ResNet-50, VGG-16, VGG-19, DenseNet, Xception, MobileNet, ResNet-152, ResNet-101 Proposed Model 92.36, 91.54, 90.36, 89.58, 93.48, 93.98, 92.78, 92.76, 94.53 respectively. The average mean score of 95.2, F1-Score of 94.01, Recall value of 94.89 and precision of 94.23. Proposed suggested model was better than all the most recent cutting-edge methods for EL picture recognition. Full comparison between suggested model as well as the most advanced models for EL picture recognition that are currently on the market. All of the existing state-of-the-art (SOTA) models were not as good as proposed models. So, the pre-trained ViT model can be used for using

EL photos to help computers find flaws in solar cells. This speeds up the process and makes sure it works every time.

VI. References

[1] F. Dincer, “The analysis on photovoltaic electricity generation status, potential and policies of the leading countries in solar energy,” Renewable and sustainable energy reviews, vol. 15, no. 1, pp. 713–720, 2011.

[2] Y. Jestin, S. Chandra, B. Cass, H. Ahmed, and S. McCormack, “Down-shifting of the incident light for photovoltaic applications,” 2021.

[3] Y.Chen,K.VanSant,Y.S.Khoo,Z.Wang,W.Luo,C.A.Deline,P.L.Hacke,J.Chai, L. Yin, Y. Wang et al., “Investigation of correlation between field performance and indoor acceleration measurements of potential induced degradation (pid) for c-si pv modules,” National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2017.

[4] T. Pei and X. Hao, “A fault detection method for photovoltaic systems based on voltage and current observation and evaluation,” Energies, vol. 12, no. 9, p. 1712, 2019.

[5] C. Schuss, K. Remes, K. Leppänen, J. Saarela, T. Fabritius, B. Eichberger, and T.Rahkonen,“Detectingdefectsinphotovoltaicpanelswith thehelpofsynchronized thermography,” IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 5, pp. 1178–1186, 2018.

[6] 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.

[7] T. Fuyuki, H. Kondo, T. Yamazaki, Y. Takahashi, and Y. Uraoka, “Photographic surveying of minority carrier diffusion length in polycrystalline silicon solar cells by electroluminescence,” Applied Physics Letters, vol. 86, no. 26, p. 262108, 2005.

[8] G. A. dos Reis Benatto, C. Mantel, S. Spataru, A. A. S. Lancia, N. Riedel, S. Thorsteinsson, P. B. Poulsen, H. Parikh, S. Forchhammer, and D. Sera, “Dronebased daylight electroluminescence imaging of pv modules,” IEEE Journal of Photovoltaics, vol. 10, no. 3, pp. 872–877, 2020.

[9] M. Mayr, M. Hoffmann, A. Maier, and V. Christlein, “Weakly supervised segmentation of cracks on solar cells using normalized l p norm,” in 2019 IEEE International Conference on Image Processing (ICIP). IEEE, 2019, pp. 1885– 1889.

[10] H. Chen, H. Zhao, D. Han, W. Liu, P. Chen, and K. Liu, “Structure-aware-based crack defect detection for multicrystalline solar cells,” Measurement, vol. 151, p. 107170, 2020.

[11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly et al.,“An image is worth 16x16 words: Transformers for image recognition at scale,” arXiv preprint arXiv:2010.11929, 2020.

[12] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” Advances in neural information processing systems, vol. 30, 2017.

[13] S. Deitsch, V. Christlein, S. Berger, et al., “Automatic classification of defective photovoltaic module cells in electroluminescence images,” Sol. Energy 185, 455–468 (2019).

[14] R. Pierdicca, E. S. Malinverni, F. Piccinini, et al., “Deep convolutional neural network for automatic detection of damaged photovoltaic cells,” in The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (Copernicus GmbH, 2018), Vol. XLII–2, pp. 893–900.

[15] W. Tang, Q. Yang, K. Xiong, et al., “Deep learning based automatic defect identification of photovoltaic module using electroluminescence images,” Sol. Energy 201, 453–460 (2020).

[16] N. V. Sridharan and V. Sugumaran, “Convolutional Neural Network based Automatic Detection of Visible Faults in a Photovoltaic Module,” Energy Sources, Part A: Recovery, Utilization, and Environmental Effects 1–16 (2021).

[17] D. Korkmaz and H. Acikgoz, “An efficient fault classification method in solar photovoltaic modules using transfer learning and multi-scale convolutional neural network,” Engineering Applications of Artificial Intelligence 113, 104959 (2022).

[18] B. Su, H. Chen, and Z. Zhou, “BAF-Detector: An Efficient CNN-Based Detector for Photovoltaic Cell Defect Detection,” IEEE Trans. Ind. Electron. 69(3), 3161–3171 (2022).

[19] X. Zhang, T. Hou, Y. Hao, et al., “Surface Defect Detection of Solar Cells Based on Multiscale Region Proposal Fusion Network,” IEEE Access 9, 62093–62101 (2021).

[20] S. Xu, H. Qian, W. Shen, et al., “Defect detection for PV Modules based on the improved YOLOv5s,” in China Automation Congress (2022), pp. 1431–1436.

[21] H. Chen, M. Song, Z. Zhang, et al., “Detection of Surface Defects in Solar Cells by Bidirectional-Path Feature Pyramid Group-Wise Attention Detector,” IEEE Trans. Instrum. Meas. 71, 1–9 (2022).

[22] Y. S. Balcıoğlu, B. Sezen, and C. Cubukcu Cerasi, “Solar Cell Busbars Surface Defect Detection Based on Deep Convolutional Neural Network,” IEEE Latin Am. Trans. 21(2), 242–250 (2023).