International Journal of Research Publication and Reviews, Vol 4, no 6, pp 1839-1846 June 2023

WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Explainable Photovoltaic Cell Defect Classification from Electroluminescence Images using Modern Deep Learning Technique

***Divyam Singh**[*,ζ], **Nalendra Namdeo**[‡]*

[ᵀ,ζ] *Aditya College of Technology and Science, M.P., India (Affiliated to RGPV, M.P., India)*

[‡] *Aditya College of Technology and Science, M.P., India (Affiliated to RGPV, M.P., India)*

A B S T R A C T

For a considerable amount of time, the upkeep of photovoltaic (PV) power plants on a large scale has been regarded as a significant obstacle. The present study introduced a defect detection approach for photovoltaic (PV) modules that employ electroluminescence images (EL) and is based on deep learning techniques. Our study presents a novel approach for automatic defect classification utilizing a pre-trained Vision Transformer deep learning model. The model is trained on EL images sourced from the publicly accessible ELPV Dataset. The model we put forth underwent training on the ELPV Dataset and achieved a test set accuracy of 98.63%. The model was subjected to rigorous experimentation, utilizing established machine learning models such as ResNet50, ResNet101, ResNet-152, VGG-16, VGG-19, DenseNet, Inception ResNet, Xception, and MobileNet as benchmarks for comparison. The results of these experiments revealed that our model outperformed all of the aforementioned models. In addition, a comparative analysis was conducted between our proposed model and the current state-of-the-art (SOTA) models for detecting image defects in EL. The results indicate that our model outperformed all existing SOTA models in terms of accuracy. The findings demonstrate that the deep learning approach presented in this study is capable of performing defect detection in electroluminescence images with high efficiency and accuracy.

Keywords: Solar Cell; Photovoltaic cell images; deep learning; explainable AI.

1. Introduction

Renewable energies are sources of clean, carbon neural and greenhouse gases free energy. The statistics produced by International Energy Agency (IEA) the demand for electricity will increase up to 70% by 2040 mainly driven by India, China, South-East Asia and Middle East. The solar energy can potentially reduce our dependence on fossil fuels significantly. The abundance of sunlight, comparatively ease of operation, cost-effectiveness and safety of operation motivates government as well as private organizations to install solar energy setup [1]. In accordance with the recent energy summit, most countries are intending to utilize solar energy as a alternative source of energy resource to ease their dependencies on the non-renewable energy such as coal energy. The Photovoltaic Cells (PV) convert photons to electrons to electricity by absorbing the light energy. PV cells are generally categorized into two classes- mono-crystalline, and poly-crystalline PV cells based on the crystal structure. The mono-crystalline cells, which are formed from a single crystal, produce comparatively more amount of electricity. On the other hand, polycrystalline PVs, which consist of multiple fragments of crystalline material, are lesser efficient, and develop faults quicker [2]. Since, MC, and PC PVs have different characteristics, periodic, automated classification and grading of the photovoltaic cells is imperative for accurate quantification and prediction of the energy in large scale solar plants. Moreover, the quality of the photovoltaic materials and the presence of faults considerably influence the electricity production. However, the PV cells tend to develop different types of faults, and cracks during the process of manufacturing, installation, transportation, or operation. These faults in turn can leads to loss of efficiency which create difficulties in load forecasting, and lead to power shortage jeopardizing the industrial sector [3]. These predicaments necessitate accurate, real-time, on-destructive detection of faults and cracks in the PV cells for efficient operation of solar power generation facilities.

The faults in the PV cells can be analysed from the electrical voltage, current and power, thermal and visual images [4]. However, the electrical attributes are generally unable to detect small and micro cracks or other defects as these small cracks do not lead to noticeable change in the current and voltage [5]. Besides, the electrical measurements cannot locate the faults, and can only be applied to individual PV cells which is not a feasible solution in practical setting. On the other hand, thermal imaging-based techniques may not be effective in all conditions as high temperature do not necessarily correspond to cracks [6]. Electro luminance imaging (EL) technique [7], which is a non-invasive method, has demonstrated its proficiency in analysing human cells. The EL imaging technology captures the images using charged coupled device around the electromagnetic wavelength range 950-1200nm in the forward bias condition. In general, the EL imaging highlights the defective portions in darker appearance that are very difficult to notice under naked or unaided eye conditions. The periodic manual inspection of the maintenance team is imperative to make critical operational decisions for this large-scale solar power generating stations. The PV cells based on their physical conditions are either retained or replaced. However, due to the remote location of such stations, periodic manual inspection is very inconvenient and costly. Hence, automated solution utilizing EL imaging is required to overcome this challenge. Moreover, the non-invasive property of EL imaging makes it convenient for automated aided solution [8]. Recent advancement in the field of Deep Learning (DL) has motivated many researchers to explore automated fault identification from EL imaging [9] [10].

In this work, we have proposed a pre-trained Vision Transformer deep learning model capable of identifying the faults in photovoltaic cells from EL images. The Vision Transformer [11] is a model architecture that is solely founded on the Transformer model [12], which has garnered significant attention in recent times due to its exceptional performance in machine translation and other natural language processing (NLP) tasks [13]. The Transformer model adheres to the encoder-decoder framework and is capable of parallel processing of sequential data without the need for any recurrent neural network. The efficacy of Transformer models has been predominantly attributed to the self-attention mechanism, which has been introduced to capture extensive interdependencies among the constituents of a sequence.

The main contributions of this paper are as follows:

- We propose a pre-trained Vision Transformer, a DL model for automated classification of EL images from the ELPV dataset.
- In the end, cross-validation and testing of the model were carried out on the images.
- The proposed pre-trained ViT model is evaluated in terms of various performance metrics, such as accuracy, F_1-score, sensitivity, and precision. And weighted scores for the same. We have also compared our proposed model with nine distinct pre-trained DL models and existing state-of-the-art.

Ongoing, the paper is structured as follows: Section II outlines the technical structure of Pre-trained ViT model. Experimental results are discussed in section III. Section IV concludes the paper.

1.1 Background and Research Motivation

Over the recent years, EL imaging-based PV fault identification approaches have become popular. In one of the initial works, Tsai et al. [14] identified the crack using the so-called independent component analysis (ICA), Spataru et al. [15] detected the presence of microcracks using matched filters. Another related work by Sovetkin et al. [16] detected the cracks in PV by performing image transformations such as the Hough transformation. Mayr et al. [9] identified the cracks and performed segmentation of the cell images by employing ResNet50 with L_p norm. Demicri et al. [17] came up with an efficient transfer learning strategy for fault detection. A related work by Parikh et al. [18] detected the presence of faults and categorized the faults from hand-crafted features, and traditional classification strategies. Chen et al. [19] explored the efficacy of steerable filtering for identifying the cracks in solar cells. Another follow-up work by Chen et al. [10] detected the presence of PV cell cracks by analysing the structural information of the images, Anwar et al. [20] explored diffusion methods for the identification of cracks in polycrystalline PV cells. Although these above-mentioned have advanced the state-of the-art, these approaches require further performance enhancement for wider practical industrial deployment. Besides, most of these works are unable to detect minor faults due to the complex pattern. Additionally, the prevalent approaches fail to perform satisfactorily in the presence of textured surface, low-contrast background, noise, and other perturbations.

The literature review determines the limitations of detecting small cracks or utilizing high computational costs with low performance. The CNNs on existing studies have performed remarkably well [21] [22] and achieved a high-performance rate. Based on excellent performance and feasibility of computational, our proposed pretrained Vision Transformer model performs with more stable performance in complex image structures such as low light and bad contrast utilizing less computational cost and more generalization strategies.

2. Material and Methodology

2.1 Dataset for Solar Photovoltaic Cell Defect Analysis

The proposed pre-trained Vision Transformer model was validated on an publicly available[1] solar cell dataset [23] [24] [25] from high resolution of 300 × 300 EL imaging from 44 PV module. There are total 2624 EL images which are categorized into two classes monocrystalline (18 module) and polycrystalline (26 module) PV module. The images were captured in a dark room to maintain uniform illumination as PV modules emits only light during the acquisition. The representative solar cell images are given in Fig.1 The extracted cell images were also randomly presented to an expert, the main focus was on defects where loss of power was > 3%. The assessments of functional and defective cells and in non-confident criteria were labelled as defective. Lower weight were assigned to control-rate's uncertainty, 33% and 67% to non-confident assessment of functional and defective cells respectively. The dataset was divided into three sets namely, training consisting of 70%, test consisting of 20%, and validation set consisting of 10% of the images. Further, the image was resized to 224 x 224 pixels for all training, testing, and validation purposes.

1 https://github.com/zae-bayern/elpv-dataset

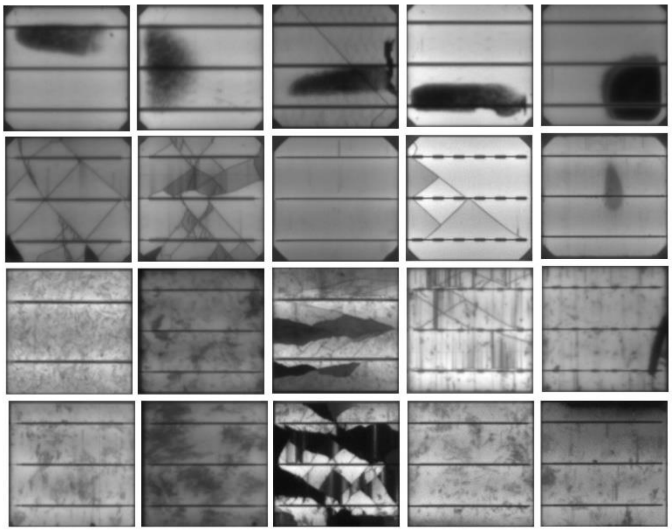

Fig. 1 Sample images from ELPV dataset; First row: Mono; second row: Poly

2.2. Proposed Pre-trained Vision Transformers Model

The Vision Transformer has been proposed as a means of expanding the applicability of the conventional Transformer to the task of image classification. The primary objective is to establish generalization across modalities beyond text while refraining from incorporating any architecture that is specific to the data. The Vision Transformer algorithm employs the Transformer's encoder module to execute classification tasks. This is achieved by mapping a sequence of image patches to their corresponding semantic label. The Vision Transformer's attention mechanism enables it to attend to various regions of an image and integrate information across the entire image, in contrast to conventional CNN architectures that typically employ filters with a local receptive field. Figure 2 illustrates the comprehensive end-to-end architecture of the model. Typically, the architecture consists of three main components: an embedding layer, an encoder, and a final head classifier. During the initial stage, a training set image X (wherein the image index i is excluded for simplicity) is partitioned into non-overlapping patches. The Transformer perceives each patch as a discrete token. In the context of processing an image X, which has a size of c x h x w (where h denotes the height, w denotes the width, and c represents the number of channels), it is common practice to extract patches of size c x p x p from the image. The aforementioned constitutes a series of patches denoted by (x_1, x_2, . . ., x_n) with a length of n, where n is equivalent to hw/p^2. The selection of patch size p is commonly either 16×16 or 32×32. It is worth noting that a smaller patch size leads to an elongated sequence, while a larger patch size results in a shorter sequence.

2.2.1. Pre-trained Vision Transformer Variants

In order to investigate the impact of augmenting the model size on classification accuracy, various iterations of the Vision Transformer (ViT) have been introduced in the work [11]. These include the "ViT-Base", "ViT-Large", and "ViT-Huge" versions. The three iterations exhibit variations in the quantity of layers employed by the encoder, the magnitude of the hidden dimension, the quantity of attention heads utilized by the MSA layer, and the size of the MLP classifier. Each of the aforementioned models undergoes training using a patch of dimensions 16×16 and 32×32. The model, denoted as "ViT-Base" comprises of an encoder that consists of 12 layers, each with a hidden size of 768. Additionally, the model employs 12 heads in the attention layer. The alternative iteration employs augmented numerical values, such as the "ViT-Large" model, which comprises 24 layers, 16 attention heads, and a hidden dimension of 1024 units. The model, denoted as "ViT-Huge" comprises 32 layers, 16 attention heads, and a hidden size of 1280.

2.2.2. Choosing the ViT-Base model

Results from experiments regarding Vision Transformers of varying sizes indicate that achieving greater accuracy necessitates the use of comparatively deeper models. Furthermore, the selection of a diminutive patch dimension leads to an augmentation in the sequence length denoted as n, thereby enhancing the model's comprehensive precision. A noteworthy discovery is that the attention heads located in the initial layers of the Vision Transformer exhibit the ability to attend to image regions that are situated at considerable distances. The proficiency in question exhibits a positive correlation with the complexity of the model. This contrasts with models based on Convolutional Neural Networks (CNN), wherein the lower layers are capable of detecting solely local information, while global information can only be detected at higher layers within the network. The aforementioned characteristic of the Vision Transformer holds significant importance in the identification of relevant features for the purpose of classification. Nevertheless, we choose ViT-Base with 80 million parameters as our model of choice.

3. Experimental Results & Discussion

3.1 Experimental Settings and Performance Evaluation

The proposed pre-trained ViT-Base model was implemented using TensorFlow and Keras libraries. Pre-trained ViT is trained on ADAM optimizer with a learning rate of 10^{-3} and a batch size of 64 for 100 epochs, and all the layers in the model are made tuneable. The dataset was divided into three sets namely, training consisting of 70 %, test consisting of 20%, and validation set consisting of 10% of the images. Further, the image was resized to 224 x 224 pixels for all training, testing, and validation purposes. The standard performance evaluation metric accuracy, and weighted and macro precision, recall, and F1 score, for consistency and multi-class EL image assessment.

3.2 Classification Results from EL Images

Table 1 describes the accuracy among other metrics obtained by our proposed pre-trained Vision Transformer model for the task of fault detection in EL images from the ELPV dataset.

Table 1 - PEM (%) scores for proposed pre-trained ViT

Accuracy	Average Type	F_1-score	Recall	Precision	Kappa
98.63	Macro	98.52	98.57	98.50	**98.40**
	Weighted	**98.64**	**98.63**	**98.66**	

3.3 Comparative Evaluation

Pre-trained networks are based on a prevalent image classification technique in the field of Deep Learning, existing state-of-the-art methods are used for comparative analysis of the proposed pretrained ViT model.

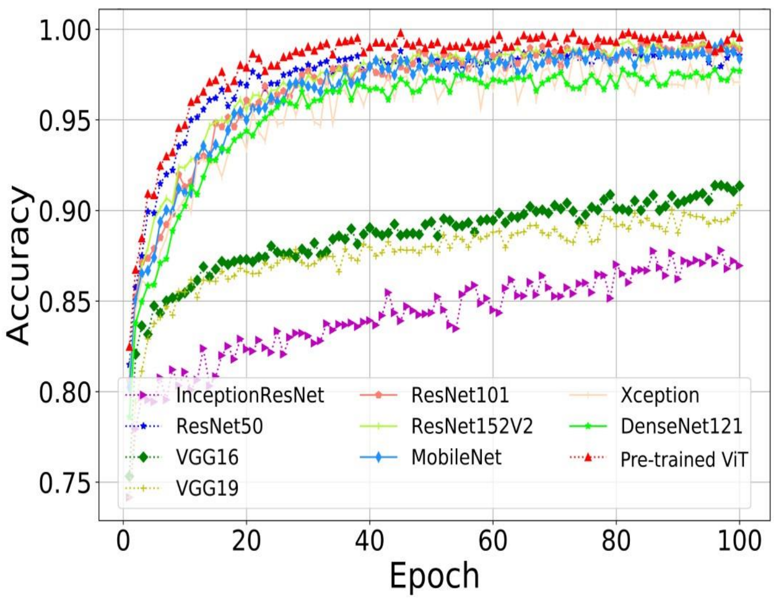

Fig. 2 Pre-trained Vision Transformer vs CNN-based pre-trained models accuracy over the epochs.

3.4 Pre-trained Model Comparison

The majority of the pretrained networks are trained on a large-scale database of more than a million images and 1000 classes [26]. As such, these pre-trained models are inherently rich in powerful features owing to training on widely used large-scale image datasets, which exhibit a broad spectrum of variability. We have fine-tuned the deeper layers in these networks to adapt new features pertinent to our classification task.

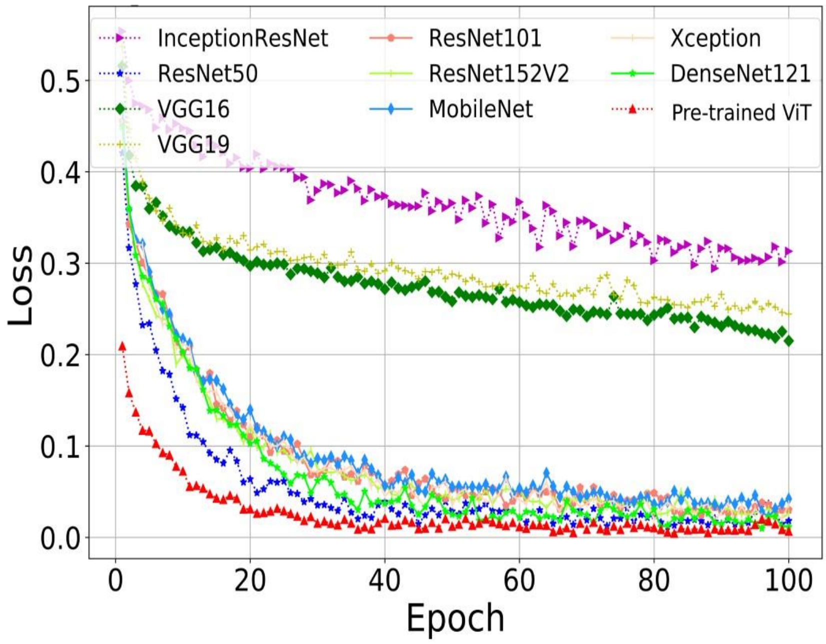

Fig. 3. Pre-trained Vision Transformer vs CNN-based pre-trained models loss over the epochs.

The variation in model accuracy and loss during the training phase for different pre-trained models on the EL dataset is shown in Fig. 2 and 3. The pre-trained ViT model reflects improved performance and superior dynamic learning than pre-trained models.

Table 2 - Result comparison with pre-trained NNs.

Network	**Depth**	**Accuracy**
ResNet-50	50	97.16
ResNet-101	101	96.76
ResNet-152	152	96.96
VGG-16	16	88.55
VGG-19	19	87.60
DenseNet	121	95.42
Inception ResNet	164	84.18
Xception	71	95.44
MobileNet	53	96.49
This work	**12**	**98.63**

The accuracy scores for the classification task on the EL dataset are tabulated in Table 2. Column 2 in Table 2 summarizes the model complexities regarding the depth and number of trainable parameters, while column 4 depicts the overall EL accuracy. The last row in Table 2 highlight pre-trained ViT's improved EL performance compared to other pre-trained networks. The pre-trained ViT architecture demonstrates the highlighted EL accuracy with a reduced number of network depth. Further, the pre-trained ViT model outperforms the ResNet-50 [27], ResNet-101 [27], ResNet-152 [27], VGG-16 [28], VGG-19 [28], DenseNet [29], Inception [30], Inception ResNet [31], Xception [32], and MobileNet [33] transfer learning models by 1.47%, 1.87%, 1.67%, 10.08%, 11.03%, 3.21%, 14.45%, 3.19%, and 2.14%, respectively.

Comparison with the Prior Art

The proposed pre-trained ViT model was also compared with existing SOTA models for the classification of ELPV dataset in table 3. Further, the pre-trained ViT model outperforms the [34], [34], [35], [36], [37], [38], [39], [40], [41], models by 16.19%, 10.21%, 5.61%, 15.63%, 10.44%, 6.93%, 4.11%, 3.23%, and 5.63%, respectively. This shows that our model performs significantly better than all the existing SOTA models.

Table 3 - State-of-the-art comparison for Classification task from EL dataset.

Ref, YoP	Method (Feature + Classifier)	Accuracy (%)
[34], 2019	KAZE, SIFT, SURF, SVM	82.44
[34], 2019	VGG19-CNN	88.42
[35], 2019	CNN	93.02
[36], 2020	GAN-CNN	83.0
[37], 2020	Hybrid Fuzzy-CNN	88.19
[38], 2021	Hybrid Attention Transformer	91.7
[39], 2021	DFB-SVM	94.52
[40], 2022	Deeplabv3+ResNet50	95.4
[41], 2022	PCA+KNN	93.0
This work	**Pre-trained ViT**	**98.63**

Conclusion

This study proposes a custom pre-trained ViT model for the classification of EL images from the ELPV dataset. Our proposed model performed better than all SOTA models for EL image classification with an average test accuracy of 98.63%. We also compared our proposed model with existing SOTA models for EL image classification, and our model performed better than all the existing SOTA models. Because of this, the pre-trained ViT model has the potential to be utilized for computer-assisted fault detection in solar cells using EL images, and it will, in turn, make the process easier and fool proof.

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,"Detectingdefectsinphotovoltaicpanelswiththehelpofsynchronized 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] J. Devlin, M.-W. Chang, and K. Lee, "Google, kt, language, ai: Bert: pre-training of deep bidirectional transformers for language understanding," in *Proceedings of NAACL-HLT*, 2019, pp. 4171–4186.

[14] 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.

[15] S. Spataru, P. Hacke, and D. Sera, "Automatic detection and evaluation of solar cell micro-cracks in electroluminescence images using matched filters," in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*. IEEE, 2016, pp. 1602–1607.

[16] E. Sovetkin and A. Steland, "Automatic processing and solar cell detection in photovoltaic electroluminescence images," *Integrated Computer-Aided Engineering*, vol. 26, no. 2, pp. 123–137, 2019.

[17] M. Demirci, N. Bes¸li, and A. Gümüs¸çü, "Defective pv cell detection using deep transfer learning and el imaging," *Proceedings Book*, p. 311, 2019.

[18] H. R. Parikh, Y. Buratti, S. Spataru, F. Villebro, G. A. D. Reis Benatto, P. B. Poulsen, S. Wendlandt, T. Kerekes, D. Sera, and Z. Hameiri, "Solar cell cracks and finger failure detection using statistical parameters of electroluminescence images and machine learning," *Applied Sciences*, vol. 10, no. 24, p. 8834, 2020.

[19] H. Chen, H. Zhao, D. Han, and K. Liu, "Accurate and robust crack detection using steerable evidence filtering in electroluminescence images of solar cells," *Optics and Lasers in Engineering*, vol. 118, pp. 22–33, 2019.

[20] S. A. Anwar and M. Z. Abdullah, "Micro-crack detection of multicrystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, pp. 1–17, 2014.

[21] M. Zhang, W. Li, R. Tao, H. Li, and Q. Du, "Information fusion for classification of hyperspectral and lidar data using ip-cnn," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022.

[22] M. Wieczorek, J. Siłka, M. Woz´niak, S. Garg, and M. M. Hassan, "Lightweight convolutional neural network model for human face detection in risk situations," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 7, pp. 4820–4829, 2022.

[23] C. Buerhop-Lutz, S. Deitsch, A. Maier, F. Gallwitz, S. Berger, B. Doll, J. Hauch, C. Camus, and C. J. Brabec, "A benchmark for visual identification of defective solar cells in electroluminescence imagery," in *35th European PV Solar Energy Conference and Exhibition*, vol. 12871289, 2018, pp. 1287–1289.

[24] 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.

[25] S. Deitsch, C. Buerhop-Lutz, E. Sovetkin, A. Steland, A. Maier, F. Gallwitz, and C. Riess, "Segmentation of photovoltaic module cells in uncalibrated electroluminescence images," *Machine vision and applications*, vol. 32, no. 4, p. 84, 2021.

[26] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *CVPR09*, 2009.

[27] 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.

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

[29] 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.

[30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V.Vanhoucke,andA.Rabinovich,"Goingdeeperwithconvolutions,"in*Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[31] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 31, no. 1, 2017.

[32] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[33] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[34] H.-N.-R. Shin, J. I. Kang, and O. Güngör, "Automatic classification of defective photovoltaic module cells in electroluminescence images," 2020.

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

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

[37] C. Ge, Z. Liu, L. Fang, H. Ling, A. Zhang, and C. Yin, “A hybrid fuzzy convolutionalneuralnetworkbasedmechanismforphotovoltaiccelldefectdetection with electroluminescence images,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 32, no. 7, pp. 1653–1664, 2021.

[38] X. T. Xie, H. Liu, Z. Na, X. Luo, D. Wang, and B. Leng, “Dpit: Detecting defects of photovoltaic solar cells with image transformers,” *IEEE Access*, vol. 9, pp.154292–154303, 2021.

[39] M. Y. Demirci, N. Besli, and A. Gümüs¸çü, “Efficient deep feature extraction and classification for identifying defective photovoltaic module cells in electroluminescence images,” *Expert Syst. Appl.*, vol. 175, p. 114810, 2021.

[40] J. Fioresi, D. J. Colvin, R. Frota, R. Gupta, M. Li, H. Seigneur, S. Vyas, S. Oliveira, M. Shah, and K. O. Davis, “Automated defect detection and localization in photovoltaic cells using semantic segmentation of electroluminescence images,” *IEEE Journal of Photovoltaics*, vol. 12, pp. 53–61, 2022.

[41] S. Bordihn, A. Fladung, J. Schlipf, and M. Köntges, “Machine learning based identification and classification of field-operation caused solar panel failures observed in electroluminescence images,” *IEEE Journal of Photovoltaics*, vol. 12, pp. 827–832, 2022