Defect recognition of solar panel in EfficientNet-B3 network based on CBAM attention mechanism

Hanran Zhang *
College of Electronics and Information Engineering, Liaoning Technical University, Huludao, 125100, China
3417856993@qq.com

Zonglin Yang
School of Management, Guangdong University of Science and Technology, Dongguan, 523000, China
yangzonglin@gdust.edu.cn

Nuo Lei
Faculty of Science, Liaoning Technical University, Fuxin, 125100, China
2631650973@qq.com

ABSTRACT

Defect recognition in solar panels is critical to safeguard their performance and efficiency. Traditional image recognition models have limitations in fine-grained defect feature extraction, which affects the accuracy and efficiency of recognition. In this paper, we propose an EfficientNet-B3 network optimization model based on the CBAM attention mechanism, which significantly improves the recognition of tiny defects in solar panels by combining deep learning techniques and attention mechanisms. Experimental results show that our model exhibits high accuracy on both training and validation sets with gradually decreasing loss. The model achieves an accuracy of 95.22% in complex and variable defect categories, which is significantly better than existing baseline models. An in-depth performance evaluation shows that the model has significant advantages in key performance metrics such as precision, recall, and F1 value, demonstrating its effectiveness and adaptability in the solar panel defect recognition task.

CCS CONCEPTS

• **Computing methodologies** → Artificial intelligence; Computer vision.

KEYWORDS

EfficientNet-B3, CBAM Attentional Mechanisms, Keyword number 3, Identification of defects in solar panels

ACM Reference Format:

Hanran Zhang *, Zonglin Yang, and Nuo Lei. 2024. Defect recognition of solar panel in EfficientNet-B3 network based on CBAM attention mechanism. In *2024 International Conference on Generative Artificial Intelligence and Information Security (GAIIS 2024), May 10–12, 2024, Kuala Lumpur, Malaysia.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3665348.3665407

1 INTRODUCTION

Solar energy, as a clean and renewable energy source, occupies an increasingly important position in the global energy mix. As solar panels are widely deployed around the world, the assurance of their operational efficiency and reliability has become a focus of research and industrial attention. Defects that may occur in panels during operation, such as hot spots, cracks and dust, not only reduce the energy conversion efficiency, but also may cause safety issues [1]. Therefore, it is of great significance to develop an efficient and accurate defect recognition method for solar anels.

The traditional image detection methods for defect detection mainly include electroluminescence image detection method, infrared image detection method and photoluminescence image detection method. [2] proposed a method for detecting and localizing solar panel damage using thermography, which detects and localizes surface hot spots through image processing techniques to indicate damage or defects. [3] proposed a method for automated visual inspection of solar cell images using improved morphological and edge detection algorithms using multiple morphological and refined edge detection and tuned parameters to extract and highlight defects on solar cells. [4] used PCA and ICA methods to detect defects in solar panels through thermal imaging. However, these methods are designed based on the features of a specific defective region, which has limited adaptability and usually needs to be redesigned for specific problems, resulting in their weak generalization ability. In addition, the soiling problem faced by PV panels in outdoor environments makes it difficult for traditional algorithms to distinguish between actual damage and soiling [5]. If dirt is incorrectly recognized as damage, it will not only increase the workload of maintenance personnel, but also may lead to a large waste of resources.

In recent years, the field of deep learning has gained widespread acclaim for its exceptional capability in identifying patterns and classifying them, finding application across a myriad of areas including self-driving cars, robotics, and medical diagnostics. The scholarly work has introduced a variety of automatic classification techniques grounded in deep learning, including notable frameworks like AlexNet [6],, VGG [7], GoogleNet [8], and ResNet [9].Deitsch and colleagues introduced a pair of automated defect detection systems leveraging Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs) [10]. Similarly, Imad Zyout implemented a transfer learning-based deep convolutional neural network to assess the surfaces of photovoltaic panels for defect detection [11]. Despite their effectiveness, these approaches often require significant computational resources for large dataset analysis and may struggle with identifying uncommon or new types of defects. Consequently, this paper selects the EfficientNet-B3 architecture enhanced with the CBAM attention mechanism. This choice allows for a balanced expansion in network depth, width,

GAIIS 2024, May 10–12, 2024, Kuala Lumpur, Malaysia

ACM ISBN 979-8-4007-0956-2/24/05
https://doi.org/10.1145/3665348.3665407

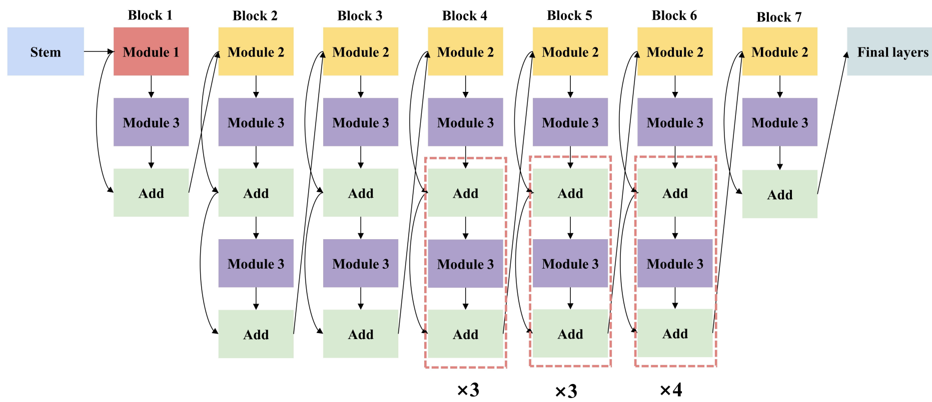

Figure 1: EfficientNet-B3 structure.

and resolution through an ingeniously simple composite coefficient, enabling better accuracy and operational efficiency without the burden of increased model size or computational demands.

2 EFFICIENTNET-B3 NET BASED ON CBAM ATTENTION MECHANISM

2.1 EfficientNet-B3 net

In the field of solar panel defect recognition, accurate extraction of panel feature information is crucial for improving recognition accuracy. Similar to the traditional image recognition task, this process not only needs to focus on the basic attributes of the panel such as color and texture, but also should take into account the microscopic defect features such as cracks, stains, etc., which are of great significance for accurately identifying various types of defects.

The EfficientNet model achieves excellent image recognition performance with low training resource consumption by adopting an efficient network structure design. One of the key innovations of the model is the introduction of the residual neural network design concept as a way to increase the depth of the network. The application of this deep network allows for more detailed and multi-layered feature extraction, thus capturing richer information in the image. By dynamically adjusting the number of feature extractors in different layers, EfficientNet is able to realize a more efficient information extraction mechanism. In addition, the model further enhances the network's ability to learn and express image details by increasing the resolution of the input image, which is particularly important for improving the accuracy of the model. In this study, EfficientNet-B3 is adopted as the basic framework. Its structure is shown in Figure 1.

2.2 EfficientNet-B3 net

The Convolutional Block Attention Module (CBAM) stands out as an effective and elegantly organized module for attention mechanisms, crafted for easy integration with any Convolutional Neural Network (CNN) setup. It synergizes two distinct sub-modules—Channel Attention and Spatial Attention—to delicately modify feature attributes across both channel and spatial realms. Initially,

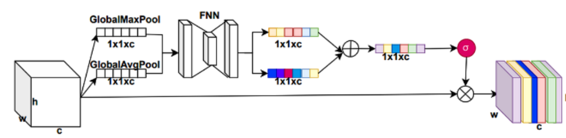

Figure 2: Channel attention module of CBAM.

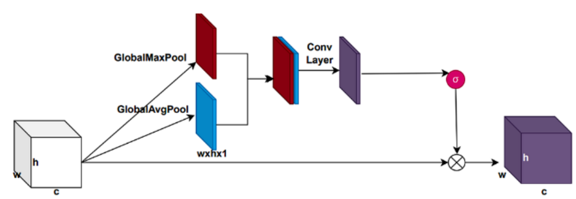

Figure 3: Spatial attention module of CBAM.

the module applies the Channel Attention sub-module to the incoming feature maps, enhancing them within the channel domain. These enhanced maps are then refined by the Spatial Attention sub-module, which further processes them to produce the final, spatially-weighted feature maps. CBAM's strategy significantly boosts the network's capacity to analyze input features by highlighting crucial data and minimizing the less relevant information. The workings of CBAM, including its channel and spatial attention mechanisms, are illustrated in Figure 2 and Figure 3, demonstrating its dual-modulation capability.

2.3 Grad-CAM

Grad-CAM is a visual interpretation method used in deep learning models, especially in CNN. It generates a rough localization map by using the gradient information of the target concepts to highlight the regions of the image that are important for the predicted concepts.

The core idea of Grad-CAM is to utilize the feature maps of the last convolutional layer of the CNN and the gradient information of the target category. Specifically, for a given category c, Grad-CAM first computes the gradient of the category c for each feature map, and then performs Global Average Pooling (GAP) on these gradients to obtain the weights w_k^c for each feature map. These weights represent the importance of each feature map for category c. Finally, all feature maps are multiplied and summed with their corresponding weights to obtain the class activation map, i.e., Grad-CAM.

$$L^c_{\mathrm{Grad-CAM}} = \mathrm{ReLU}\left(\sum_k w_k^c A^k\right) \tag{1}$$

where A^k is the kth feature map and $\mathrm{I}_{\mathrm{Grad-CAM}^c}$ is the class activation map for category c. The ReLU function is used to retain only positive influences, i.e., to focus only on those features that have a positive predictive effect on category c. With Grad-CAM, we can visualize the regions that the model focuses on when making decisions, providing some interpretability to the model.

3 EXPERIMENTAL AND DATA RESULTS

3.1 Experimental environment setup

In this paper, we use python 3.6.8 language and PyTorch 1.10 framework for deep learning experiments, and use CUDA version 11.4 and NVIDIA RTX3050 graphics card to accelerate the computation. In the experiments, we set the learning rate to 0.001, the batch size to 32, the epoch number to 100, and the optimizer to Adam. We use cross entropy as the loss function, and accuracy, precision, recall and F1 score as the evaluation metrics.

To ensure the reproducibility of the experimental results, we rigorously configured and tested the experimental environment. We also conducted experiments with different hyperparameter settings to assess their impact on model performance. In addition, we used a number of regularization techniques, such as Dropout and Batch Normalization, to prevent model overfitting.

3.2 Datasets

The dataset utilizes a private dataset consisting of solar panels extracted using Electroluminescence (EL) technology on a solar panel production line in Shenyang, China. Electroluminescence is an advanced inspection method that captures an image of the resulting light radiation by applying an electric current to a solar panel to demonstrate defects or irregularities that may not be visible under normal conditions, and in this paper the collected data is categorized into defective and non-defective categories.

In this paper, we perform detailed preprocessing of the dataset including denoising, contrast enhancement and image segmentation to improve the accuracy of defect detection. In addition, we employ data enhancement techniques such as rotation, flipping and scaling to increase the generalization ability of the model.

3.3 Evaluating indicator

To evaluate the accuracy of our model, we utilize four principal metrics for classifying results: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The labels "true" and "false" signify if the model's classification of images is correct or incorrect, respectively. "True positive" and "true negative" outcomes indicate accurate categorization of images as either positive or negative. In contrast, a false positive refers to an instance where the model erroneously identifies a negative observation as positive. On the flip side, a false negative is when the model mistakenly labels a positive observation as negative. By applying these four categories, we proceed to compute the following performance indicators:

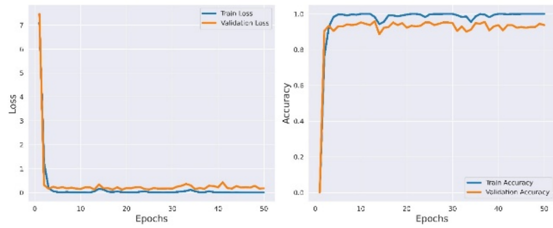

Figure 4: Training Accuracy vs. Loss Iteration Change Plot.

Table 1: Performance Comparison of Defect Recognition Models.

Model	Acc	Precision	Recall	F_1
resnet50	0.9154	0.9193	0.9154	0.9152
densenet169	0.9485	0.9501	0.9485	0.9485
mobilenet_v2	0.9154	0.9193	0.9154	0.9152
squeezenet1_0	0.9338	0.9362	0.9338	0.9337
EfficientNet-B3	0.9449	0.9460	0.9449	0.9448
EfficientNet-B3+CBAM	0.9522	0.9524	0.9522	0.9522

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

3.4 Results

To enhance the precision and speed of recognition, this study trained the model across 50 distinct batches utilizing the training dataset. The outcomes from this training indicate that the model reached a peak accuracy of 97.25% on the validation set. Figure 4 displays the model's iterative outcomes, illustrating a consistent decline in the model's loss function as the training iterations progress, alongside a corresponding improvement in localization accuracy. This trend underscores the model's capability to efficiently learn and identify defects.

After completing 50 epochs of training, the model demonstrated excellent performance on the test set with an accuracy of 95.22%. In order to fully evaluate the efficacy of the model, this paper trains and tests several other popular deep learning models under the same hardware configuration and software environment. The results are shown in Table 1.

The results show that the EfficientNet-B3+CBAM model outperforms the EfficientNet-B3 model alone in all evaluation metrics. This indicates that the addition of CBAM attention mechanism effectively improves the model's ability to recognize defective features.

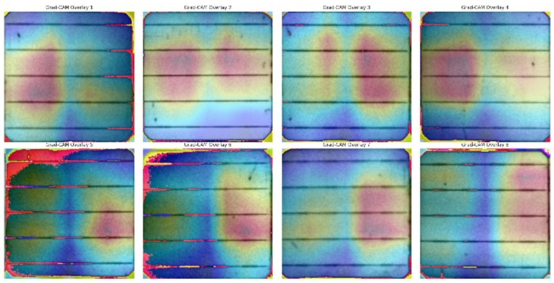

Figure 5: Grad-CAM Positioning Chart.

Meanwhile, compared with other baseline models, EfficientNet-B3+CBAM performs particularly well in terms of accuracy and F_1 scores. For example, compared with ResNet-50, EfficientNet-B3+CBAM has about 7% higher accuracy and similar improvement in F_1 scores. This result further confirms the efficiency of the EfficientNet architecture in handling image recognition tasks and the effectiveness of the CBAM module in enhancing model attention.

We used Grad-CAM to view the region of the model's attention to the image, as shown in Figure 5.

It is observed through the localization map that the model mainly focuses on the regions to the left and right of the center, as well as the edge portion of the image while processing the image. This finding indicates that the model assigns higher importance to these regions during the recognition process. The left-of-center and right-of-center focus may represent the model's search for specific defective features, such as cracks or stains, that are commonly found within these regions of the solar panel image. The heat maps generated by Grad-CAM reveal the significant role of these regions in the model's predictions; the focus on the edges of the image, on the other hand, may be due to the noise generated by the dataset during the segmentation process. During the segmentation and preprocessing stages of the solar panel images, the edge regions may have introduced noise due to mishandling, and these noises were mistakenly recognized by the model as potential defective features.

4 CONCLUSION

In this paper, an EfficientNet-B3 network based on the CBAM attention mechanism is proposed, which is specifically optimized for the solar panel defect recognition task. By combining the deep learning technique and the attention mechanism, the model is able to accurately capture the subtle defect features in the panel image, realizing the efficient recognition of panel defects.

Experimental results show that the model achieves a high accuracy of 95.22% on the solar panel defect recognition task, significantly outperforming the existing baseline model. Through in-depth performance evaluation, this paper finds that the model performs well in key performance metrics such as precision, recall, and F_1 value, which demonstrates the effectiveness of the CBAM attention mechanism in enhancing the model's recognition capability. Future work will be devoted to further improve the generalization ability of the model, especially the recognition accuracy of different types of defects in complex environments.

REFERENCES

[1] S. Prabhakaran, R. Annie Uthra, and J. Preetharoselyn, "Comprehensive Analysis of Defect Detection Through Image Processing and Machine Learning for Photovoltaic Panels," in Computer Vision and Machine Intelligence Paradigms for SDGs, R. J. Kannan, S. M. Thampi, and S.-H. Wang, Eds., Singapore: Springer Nature, 2023, pp. 245–261.

[2] S. Jaybhaye, O. Thakur, R. Yardi, V. Raut, and A. Raut, "Solar Panel Damage Detection and Localization of Thermal Images," J Fail. Anal. and Preven., vol. 23, no. 5, pp. 1980–1990, Oct. 2023.

[3] M. Cao, L. Ma, and R. Li, "Automated visual inspection of solar cell images using adapted morphological and edge detection algorithms," J Opt, Jul. 2023.

[4] J. Jia, R. Lei, L. Qin, and X. Wei, "i5mC-DCGA: an improved hybrid network framework based on the CBAM attention mechanism for identifying promoter 5mC sites," BMC Genomics, vol. 25, no. 1, p. 242, Mar. 2024.

[5] Khan, Farhan Ali, Brijesh Kumar Gautam, and Sanjay Singh. "Detection of defects in solar panels using thermal imaging by PCA and ICA method." International Research Journal of Engineering and Technology (IRJET) 4.06 2017: 1700-1703.

[6] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," Ecological Informatics, vol. 61, p. 101182, Mar. 2021.

[7] B. Ni, P. Zou, Q. Li, and Y. Chen, "Intelligent Defect Detection Method of Photovoltaic Modules Based on Deep Learning," presented at the 2018 International Conference on Transportation & Logistics, Information & Communication, Smart City (TLICSC 2018), Atlantis Press, Dec. 2018, pp. 167–173.

[8] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition." arXiv, Apr. 10, 2015.

[9] C. Szegedy *et al.*, "Going deeper with convolutions," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2015, pp. 1–9.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778. Accessed: Mar. 16, 2024. [Online].

[11] S. Deitsch *et al.*, "Automatic classification of defective photovoltaic module cells in electroluminescence images," Solar Energy, vol. 185, pp. 455–468, Jun. 2019.