YOLO-iCBAM: an improved YOLOv4 based on CBAM for defect detection

Conference Paper · March 2024

DOI: 10.1117/12.3023071

2 authors:

Junqi Bao
Macao Polytechnic University

Xiaochen Yuan
Macao Polytechnic University

YOLO-iCBAM: An Improved YOLOv4 based on CBAM for Defect Detection

Junqi Bao and Xiaochen Yuan*

Faculty of Applied Science, Macao Polytechnic University, Macao, China 999078

ABSTRACT

Defect detection in Photovoltaic (PV) cell Electroluminescence (EL) images is a challenge in industry. In this paper, a novel defect detection method YOLOv4 with an improved Convolutional Block Attention Module (YOLO-iCBAM) is proposed for PV cell EL images. We first propose an improved CBAM to enhance the network's ability to capture multi-scale defects in complex image backgrounds. Then, we modify the conventional YOLOv4 architecture for defect detection. Specifically, we adjust the backbone network to make a fast convergence. Then, we adopt the iCBAM to YOLOv4 to refine the feature map before YOLO Head. Then, we train a K-Means++ model based on PV cell EL images to generate anchors for bounding box regression. Moreover, we conduct experiments in the PVEL-AD dataset to evaluate the proposed YOLO-iCBAM. The experimental results indicated that the proposed YOLO-iCBAM achieves a better F1-Score of 0.716 and mAP of 0.748.

Keywords: Defect Detection, Photovoltaic Cell, Deep Learning, Computer Vision

1. INTRODUCTION

Multicrystalline photovoltaic (PV) cells are used in the field of solar energy systems to convert solar power into electricity. However, defects in the PV cells may be produced during manufacturing by accident, which will lead to an inefficient conversion and reduce the life span of PV cells.[1] Thus, computer-aided defect detection on PV cells has become a popular challenge in the industry. Since electroluminescence (EL) imaging captures the PV cells with a high-resolution,[2] it is easier for researchers to detect defects in PV cells using EL images.[3] Defect detection for PV cell EL images is to localize the region of the defect and identify the categories of the defects. In recent years, many state-of-the-art methods have been proposed, and these methods can be roughly divided into two categories: handcrafted feature-based methods and learning-based methods.

The handcrafted feature-based defect detection methods focus on extracting features from the EL images to represent the texture, structure, color, and shape of EL images. Thereafter, a data-based detector is applied to learn from the handcrafted feature and localize the defect. Tsai et al.[4] proposed an anisotropic diffusion method to detect micro-cracks, which is efficient and fast in detecting micro-cracks. However, the methods can only detect micro-cracks, which is not suitable for multi-category defects. Su et al.[5] proposed a novel feature descriptor to enhance the defect classification in the heterogeneous background. Dhimish et al.[6] adopt binary and Discrete Fourier Transform (DFT) to detect cracks in EL images. The experimental results indicated that the detection performance of micro-cracks on EL images has been improved. The handcrafted feature-based methods achieve remarkable performance in defect detection, however, these methods still have lots of limitations such as a lack of ability to detect multi-category defects.

Recently, deep learning-based methods have gained lots of attention for their advantages of wide versatility and better performance. Chen et al.[7] proposed a crack detection model based on a Convolutional Neural Network (CNN) and a Naïve Bayes (NB) data fusion scheme, called NB-CNN. In addition, Chen et al.[8] further introduced a Fully Convolutional Network (FCN) to NB-CNN and proposed a NB-FCN for crack detection. The experiments indicated that both NB-CNN and NB-FCN achieve remarkable performance in the industrial manufacturing defect classification. Su et al.[9] proposed a faster RPAN-CNN for defect detection in PV cell EL images, which achieves better performance than existing detection methods

In conventional CNN-based methods, the image is processed through convolutional and pooling layers to extract features. However, these models learn from these features equally, regardless of their different contributions to the tasks. The

* xcyuan@mpu.edu.mo

attention mechanism addresses this limitation by allowing the model to adaptively reallocate the weight of the feature map. Woo et al.[10] proposed a lightweight Convolutional Block Attention Module (CBAM), which obtains a channel attention map and a spatial attention map for feature refinement by multiplying the two attention maps. The proposed CBAM has been proven to be suitable for CNN-based models and improves the performance of classification and detection significantly.

This paper proposes a YOLOv4 model with improved CBAM (YOLO-iCBAM) for defect detection in PV cell EL images. In addition, since CBAM[10] performs unsatisfactorily when the images have a complex background, we propose an improved CBAM (iCBAM) to emphasize defect features while suppressing background noise efficiently. YOLOv4[11] is a popular CNN-based model for object detection, however, YOLOv4 is hard to detect multi-scale defects in EL images with complex backgrounds. Thus, we incorporate the proposed iCBAM to YOLOv4 to make the model detect multi-scale defects on EL images. According to the experimental results, the proposed YOLO-iCBAM achieves a better detection performance. The rest of this paper is organized as follows. Section 2 describes the detailed process of the proposed defect detection model; Section 3 conducts experiments to demonstrate the superiority of the proposed model; and Section 4 gives the conclusion and future works.

2. PROPOSED METHODS

In this section, we describe the proposed YOLO-iCBAM in detail and demonstrate the architecture in Figure 1. Specifically, we increase the number of the ResNetx in the first CSPx from 1 to 2, where x denotes the number of the ResNet. Since the positional information tends to diminish as the convolution progresses, increasing the number of CSPx helps achieve faster convergence and better performance. Secondly, we applied the proposed iCBAM on two feature maps with the sizes of 26×26 and 52×52 to refine the feature map and make the network focus on the defective features. Thirdly, we train a K-Means++[12] model based on the PV cell EL images to generate the anchors for bounding box regression.

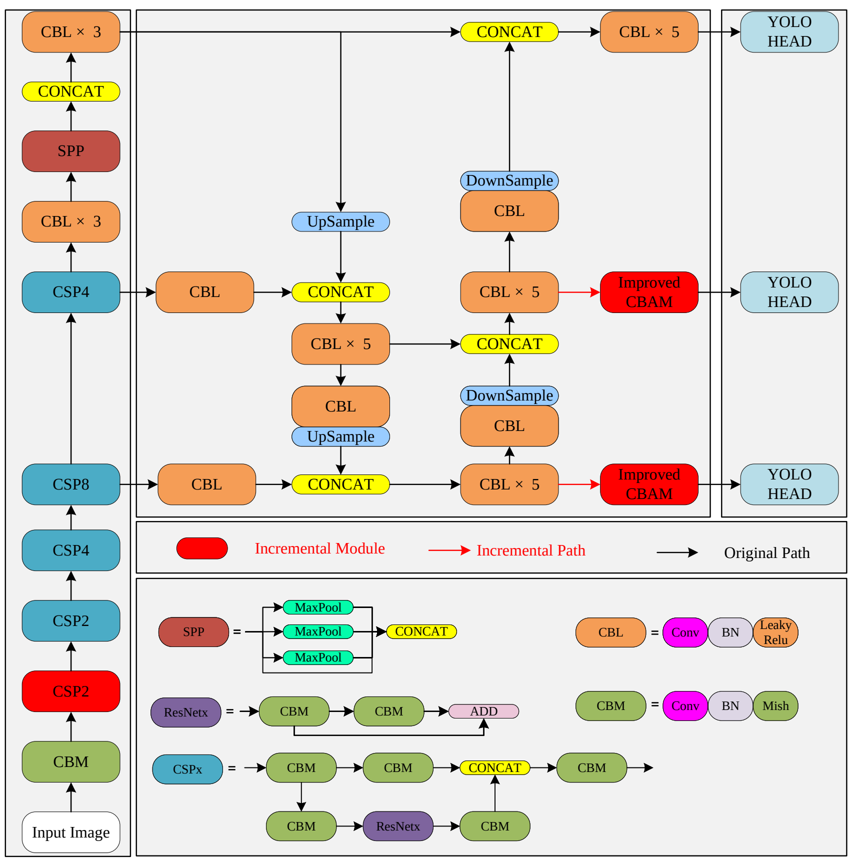

Figure 1. Architecture of the proposed YOLO-iCBAM

Inproved CBAM is proposed in this paper to help the network learn more efficient information from the channel and spatial of the features. Since CBAM[10] performs unsatisfactorily when the background of an image is complex, we propose an improved channel attention module for CBAM to enhance the anti-noise ability. Figure 2 (a) demonstrates the architecture of the proposed iCBAM. Given an input feature map $f_i \in R^{W \times H \times C}$, iCBAM calculates a channel attention map M_C and a spatial attention map M_S. The output feature map $f_o \in R^{W \times H \times C}$ is calculated according to:

$$f^{'} = M_C \otimes f_i \tag{1}$$

$$f_o = M_S \otimes f^{'} \tag{2}$$

where $\otimes$ denotes the element-wise multiplication.

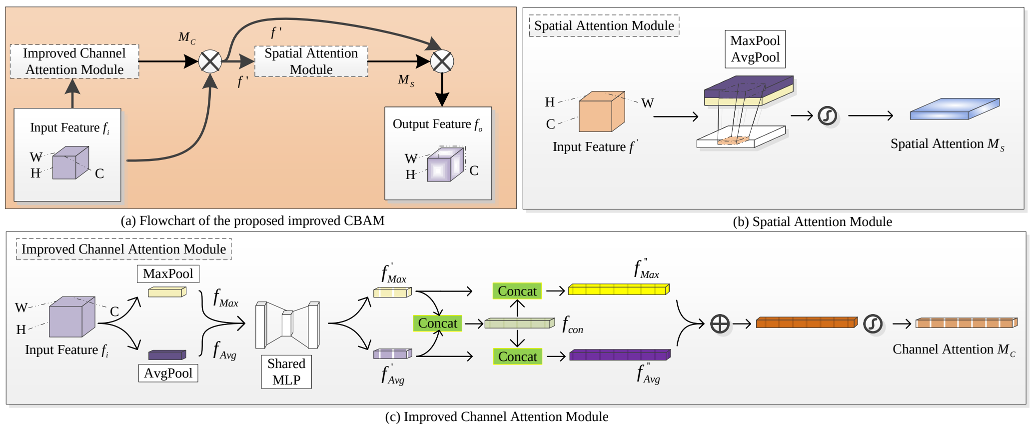

Figure 2. Architecture of the proposed iCBAM. (a) Flowchart of the proposed improved CBAM; (b) Spatial Attention Module; (c) The proposed Improved Channel Attention Module.

Figure 2 (b) demonstrates the spatial attention module. Given an input feature f^i, a convolution operation is applied to the average-pooled and max-pooled features. The spatial attention map M_S is then obtained after the activation function Sigmoid. Figure 2 (c) demonstrates the proposed Improved Channel Attention Module. The channel attention map M_C is calculated as (3). Specifically, the input feature map f_i will undergo a max pooling layer and an average pooling layer. Then, a shared Multi-Layer Perceptron is adopted to the generated f_{Max} and f_{Avg}. The output feature $f^{'}_{Max}$ and $f^{'}_{Avg}$ will be concatenated first. Thereafter, f_{con} will be merged with $f^{'}_{Max}$ and $f^{'}_{Avg}$, respectively. An element-wise summation will be applied to $f_{Max}x^{''}$ and $f^{''}_{Avg}$and a sigmoid activation function will be applied to produce the channel attention map M_C.

$$\begin{cases} f^{'}_{Max} = MLP(MP(f_i)) \\ f^{'}_{Avg} = MLP(AP(f_i)) \\ f^{''}_{Max} = Concat(f^{'}_{Max}, f^{'}_{Avg}) + f^{'}_{Max} \\ f^{''}_{Avg} = Concat(f^{'}_{Max}, f^{'}_{Avg}) + f^{'}_{Avg} \\ M_C = Sigmoid(Concat(f^{''}_{Max}, f^{''}_{Avg})) \end{cases} \tag{3}$$

3. EXPERIMENTAL RESULTS

In this section, we conduct experiments to evaluate the proposed iCBAM and YOLO-iCBAM using PVEL-AD[13] dataset. The distribution of the training, validation, and test set is demonstrated in Table 1. In addition, to evaluate the proposed YOLO-iCBAM, four evaluation metrics are used, which are listed as follows:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

$$mAP = \frac{1}{N} \sum_{0}^{N} \int_{0}^{recall} precision(recall)d(recall) \tag{7}$$

where TP denotes the true positive, FP denotes the false positive, and FN denotes. N is the number of the categories, and in this paper, $N = 5$.

Table 1. Distribution of the Training, Validation, and Test Set.

Category	Defect Number			Total
	Training	Validation	Test	
Black Core	410	39	45	404
Finger	458	79	60	597
Thick Line	417	38	57	512
Star Crack	50	9	11	70
Horizontal Dislocation	405	51	511	507
Total	1740	216	224	2180

3.1 Ablation Study

In this section, we conduct an ablation study to discuss the superiority of the proposed iCBAM. Table 2 demonstrates the comparison of detection performance between YOLOv4 using CBAM and iCBAM respectively in terms of Precision, Recall, F1-Score, and mAP. The best results have been highlighted in bold. The proposed iCBAM adopted in YOLOv4 achieves the best F1-Score of 0.716 and mAP of 0.748. CBAM[10] significantly improves the performance of classifying the defects which is reflected in Precision. However, due to the complex background in EL images, the recall has reduced compared to conventional YOLOv4. Although the proposed iCBAM only achieves the second-best in Precision of 0.645 and Recall of 0.816, it well balance the Precision and Recall and achieves the best F1-Score and mAP in defect detection. Thus, the proposed iCBAM outperforms the conventional CBAM, and can achieve a better detection performance.

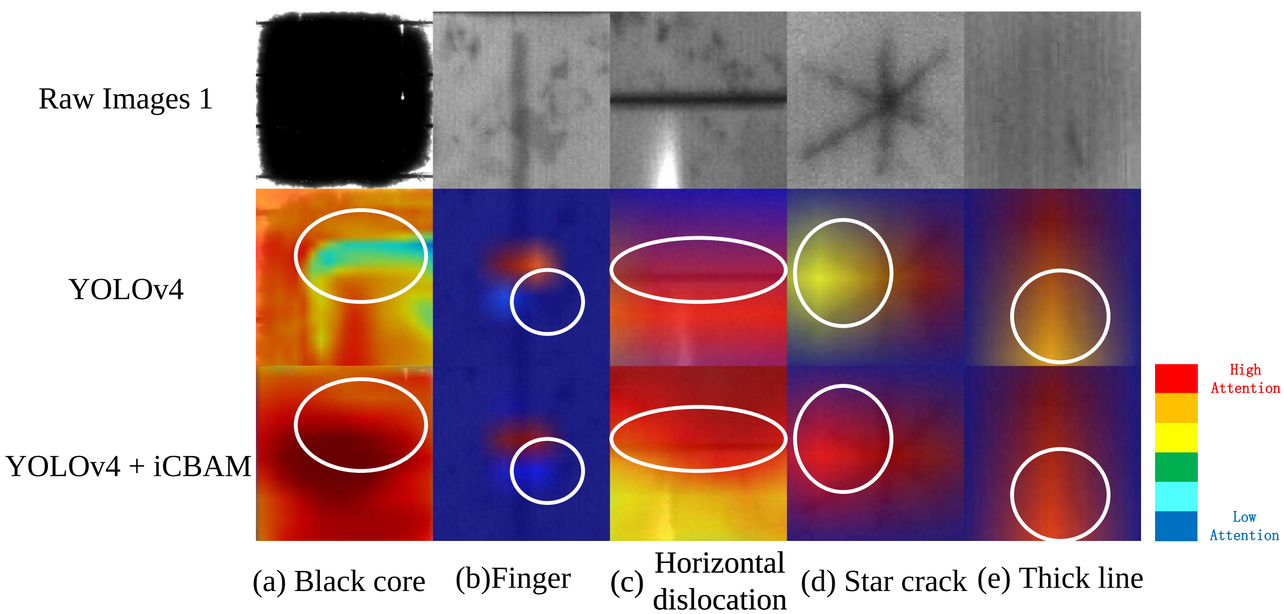

Figure 3. Demonstration of attention level of YOLOv4 with and without iCBAM. 1st row is the original raw EL images; 2nd is the attention level of 5 categories of defect using YOLOv4; 3rd is the attention level of 5 categories of defect using YOLOv4 with iCBAM

Table 2. Comparison of detection performance between YOLOv4 using CBAM and iCBAM.

Model	Group		Evaluation Metrices			
	CBAM	iCBAM	Precision	Recall	F1-Score	mAP
YOLOv4			0.622	**0.834**	0.706	0.740
YOLOv4 + CBAM	✓		**0.676**	0.766	0.712	0.712
YOLOv4 + iCBAM		✓	0.645	0.816	**0.716**	**0.748**

Figure 3 demonstrates the attention level of YOLOv4 with and without iCBAM. 1st row is the original raw EL images; 2nd is the attention level of 5 categories of defect using YOLOv4; 3rd is the attention level of 5 categories of defect using YOLOv4 with iCBAM. As shown in this figure, the proposed iCBAM can be effective in helping networks focus on defects, which is reflected in the darker color of the defective area.

3.2 Comparison with state-of-the-art methods

In this section, we conduct experiments to compare the proposed YOLO-iCBAM with several state-of-the-art methods such as Faster R-CNN,[14] YOLOv4.[11] Table 3 demonstrates the comparison of detection performance in terms of Precision, Recall, F1-Score, and mAP. The best results have been highlighted in bold. The proposed YOLO-iCBAM achieves the best Precision of 0.645, F1-Score of 0.716 and mAP of 0.748. Although conventional YOLOv4 achieves the best Recall of 0.834, its Precision is 0.023 lower than the proposed YOLO-iCBAM while the Recall is only 0.018 higher.

Table 3. Comparison of detection performance between YOLO-iCBAM and state-of-the-art methods.

Model	Evaluation Metrices			
	Precision	Recall	F1-Score	mAP
Faster R-CNN	0.483	0.581	0.560	0.658
YOLOv4	0.622	**0.834**	0.706	0.740
proposed YOLO-iCBAM	**0.645**	0.816	**0.716**	**0.748**

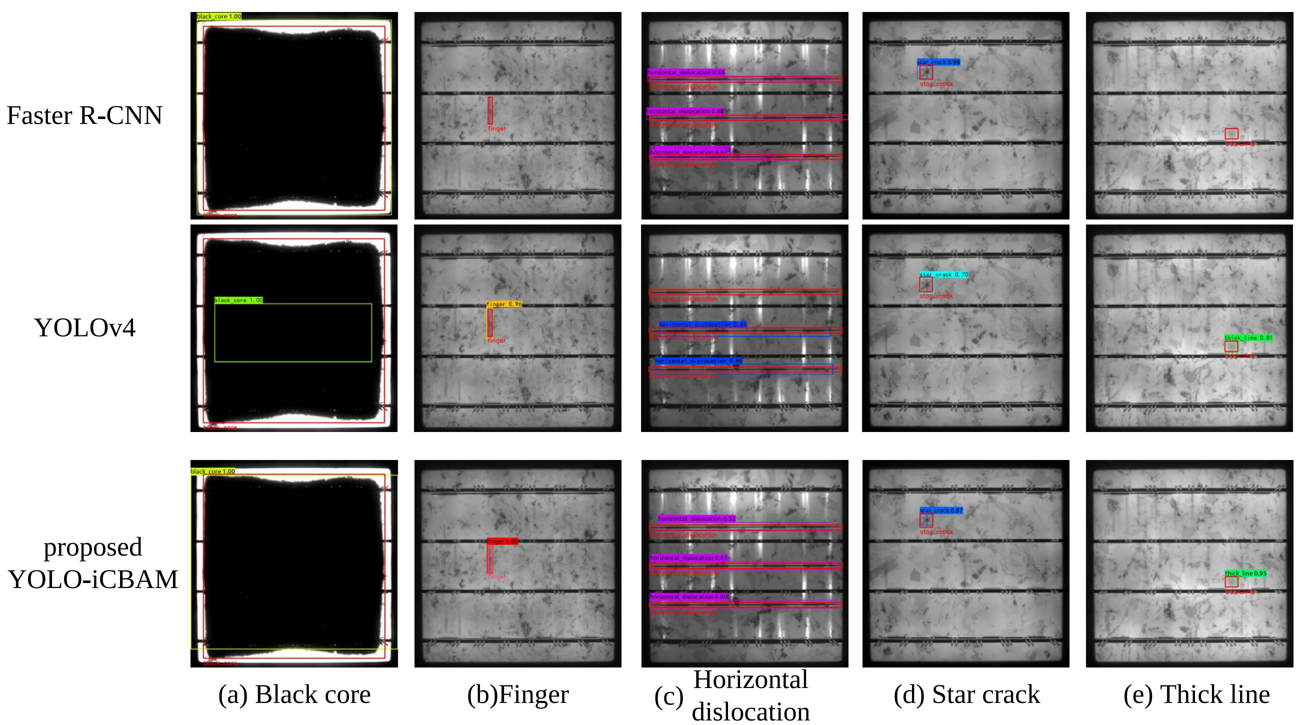

Figure 4. Demonstration of results of defect detection in PVEL-AD dataset using Faster R-CNN, YOLOv4, and proposed YOLO-iCBAM. (a) is the detection result of the Black Core; (b) is the detection result of the Finger; (c) is the detection result of the Horizontal dislocation; (d) is the detection result of the Star Crack, and (e) is the detection result of the Thick Line.

Figure 4 demonstrates the results of defect detection in PVEL-AD dataset using Faster R-CNN, YOLOv4, and the proposed YOLO-iCBAM. The red box with categories demonstrates the ground truth while the colorful box with categories and score is the predicted bounding box. As illustrated in Figure 3, Faster R-CNN and YOLOv4 have missed detection, such as Finger and Thick Line on Faster R-CNN, and Horizontal Dislocation on YOLOv4. Therefore, the proposed YOLO-iCBAM can perform better in defect detection.

4. CONCLUSION

In this paper, we propose a defect detection method YOLO-iCBAM for PV cell EL images. We first proposed an improved CBAM to allow the network to capture the defect features from a complex background more efficiently. Then we adopt the improved CBAM to YOLOv4 to enhance the ability of YOLOv4 for multi-scale defect detection. Further, we use K-Means++ to generate the anchors for bounding box regression. The experimental results have indicated that the proposed YOLO-iCBAM achieves a better detection performance for defect detection in PV cell EL images. In the future, we will focus on making the network more robust to multi-scale defects and meet the real-time requirements in the industry.

5. ACKNOWLEDGMENTS

This work was funded by the Science and Technology Development Fund of Macau SAR (Grant number 0045/2022/A), and the Macao Polytechnic University (Project No. RP/FCA-12/2022).

REFERENCES

[1] Herraiz, A. H., Marugan, A. P., and Marquez, F. P. G., "Photovoltaic plant condition monitoring using thermal images analysis by convolutional neural network-based structure," **153**, 334–348.

[2] Deitsch, S., Christlein, V., Berger, S., Buerhop-Lutz, C., Maier, A., Gallwitz, F., and Riess, C., "Automatic Classification of Defective Photovoltaic Module Cells in Electroluminescence Images," **185**, 455–468.

[3] Breitenstein, O., Bauer, J., Bothe, K., Hinken, D., Müller, J., Kwapil, W., Schubert, M. C., and Warta, W., "Can Luminescence Imaging Replace Lock-in Thermography on Solar Cells?," **1**(2), 159–167.

[4] Tsai, D.-M., Chang, C.-C., and Chao, S.-M., "Micro-crack inspection in heterogeneously textured solar wafers using anisotropic diffusion,"

[5] Su, B., Chen, H., Zhu, Y., Liu, W., and Liu, K., "Classification of Manufacturing Defects in Multicrystalline Solar Cells With Novel Feature Descriptor," **68**(12), 4675–4688.

[6] Dhimish, M. and Holmes, V., "Solar cells micro crack detection technique using state-of-the-art electroluminescence imaging," **4**(4), 499–508.

[7] Chen, F.-C. and Jahanshahi, M. R., "NB-CNN: Deep Learning-Based Crack Detection Using Convolutional Neural Network and Naïve Bayes Data Fusion," **65**(5), 4392–4400.

[8] Chen, F.-C. and Jahanshahi, M. R., "NB-FCN: Real-Time Accurate Crack Detection in Inspection Videos Using Deep Fully Convolutional Network and Parametric Data Fusion," **69**(8), 5325–5334.

[9] Su, B., Chen, H., Chen, P., Bian, G., Liu, K., and Liu, W., "Deep Learning-Based Solar-Cell Manufacturing Defect Detection With Complementary Attention Network," **17**(6), 4084–4095.

[10] Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S., "CBAM: Convolutional Block Attention Module," in [*Computer Vision – ECCV 2018*], Ferrari, V., Hebert, M., Sminchisescu, C., and Weiss, Y., eds., **11211**, 3–19, Springer International Publishing.

[11] Bochkovskiy, A., Wang, C.-Y., and Liao, H.-Y. M., "YOLOv4: Optimal Speed and Accuracy of Object Detection."

[12] Arthur, D. and Vassilvitskii, S., "K-means++ the advantages of careful seeding," in [*Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*], 1027–1035.

[13] Su, B., Zhou, Z., and Chen, H., "PVEL-AD: A Large-Scale Open-World Dataset for Photovoltaic Cell Anomaly Detection," **19**(1), 404–413.

[14] Ren, S., He, K., Girshick, R., and Sun, J., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks."