Design of photovoltaic hot spot detection system based on deep learning

Yifeng Ren, Yongjun Yu, Jing Li, Wenhua Zhang
School of Electrical and Control Engineering, North University of China, Taiyuan, 030051, China
renyifeng126@126.com, renyifeng@nuc.edu.cn

Abstract. At present, it is difficult to detect the photovoltaic (PV) hot spots and the recognition efficiency is low. In this paper, an improved Single Shot MultiBox Detector (SSD) algorithm was designed for PV hot spot detection. The algorithm used the MobileNet network to replace the VGG16 convolutional neural network structure in the original SSD. This network is a depthwise separable convolution structure. Using it for feature extraction can reduce the number of parameters in the structure and achieve the purpose of speeding up the network. The experimental results show that the improved algorithm can detect the hot spots of PV array with good confidence, low detection rate and good robustness. Compared with the You Only Look Once (YOLO) algorithm and the original SSD algorithm, the detection speed is significantly improved, which verifies the effectiveness of the algorithm.

1. Introduction
With the large-scale expansion of industry, population increase and the continuous growth of per capita energy consumption, the energy demand of almost all countries in the world is rising [1]. As a clean and sustainable energy source [2-3], PV power generation is considered to be the most reliable choice among all renewable energy sources. However, in some environments, the PV array will be blocked by trees, neighboring buildings or bird droppings. The shadow of the PV array can cause the battery to overheat, damage the PV array, and even lead to direct scrap [4]. Therefore, the hot spot detection of PV array is particularly important.

Aline and her team [5] proposed a fault inspection by using aerial infrared thermography in a PV plant. They used infrared thermography and Unmanned Aerial Vehicles to detect the faults. However, whether it is a hot spot, open circuit, short circuit or damaged module, their method requires technicians to evaluate, locate and classify the fault. Jiang Lin [6] proposed a B-spline least squares fitting method based on gray histogram for hot spot detection. The use of infrared thermal image processing improves the accuracy of detecting hot spots, but the hot spot fault points are not located. With the rise of artificial intelligence, computer vision has been applied to the field of fault detection [7]. Redmon J [8-9] and others proposed a YOLO algorithm. In the aspect of target detection, the classifier is reused to perform detection, which improves the detection speed and accuracy greatly. And [10] proposed a SSD algorithm innovatively, which discretized the output space of the bounding box into a set of default boxes. These default boxes have different aspect ratios. It is faster than YOLO algorithm. Compared with Fast-RCNN, it also has similar accuracy and more faster speed.

In this paper, an improved SSD algorithm was designed. The lightweight MobileNet network was used to replace the VGG16 convolutional neural network structure, which can reduce the number of parameters in the structure and increase the network speed. The experimental results show that the
proposed algorithm has a great improvement on the detection speed of PV module hot spots, the system robustness and confidence. Second 2 introduces a network structure and principle of the SSD algorithm. Second 3 introduces how to use the MobileNet network to improve the SSD algorithm. Second 4 discusses the experiment and results. Section 5 presents conclusions.

2. SSD network architecture

Unlike target detection models based on the region proposal such as Fast-RCNN and Faster-RCNN, SSD is based on multi-border regression. SSD is a target detection model based on bounding boxes regression. The SSD network not only uses multi-scale feature maps for detection, but also uses convolution to perform feature detection on different feature maps directly. Fast-RCNN [11] needs two steps to complete the detection. First, obtain candidate frames through CNN. Second, perform the classification and regression. However, SSD only needs one step to complete the detection. In addition, YOLO needs to be detected after the fully connected layer, while SSD uses the convolutional layer for detection directly. Its structure is shown in Fig. 1.

The input image size of the SSD model is mainly 300×300. The image enters the VGG-16 convolutional network through the input layer for feature extraction, and finally calculates the loss through the output layer. The loss function consists of two parts, called the positioning loss and the regression loss. However, the SSD network uses both large feature and small feature maps for detection, which will generate a large number of parameters and make the running speed slow.

The loss function of SSD is as formula (1):

$$
L(x,c,l,g) = \frac{1}{N} (L_{conf}(x,c) + \alpha L_{loc}(x,l,g))
$$

(1)

Where $N$ is the number of positive samples of the prior box. $L$ is the position prediction value of the corresponding bounding box of the prior box.

The position error adopts $\text{Smooth}_{L1}$ Loss, as shown in formula (2):

$$
\text{smooth}_{L1}(x) = \begin{cases} 
0.5x^2 & \text{if } |x| < 1 \\
|x| - 0.5 & \text{otherwise}
\end{cases}
$$

(2)

For the confidence error, it uses softmax loss as shown in formula (3):

$$
L_{conf}(x,c) = -\sum_{i=\text{pos}}^N x_i^p \log(\hat{c}_i^p) - \sum_{i=\text{neg}} \log(\hat{c}_i^p)
$$

(3)

To further normalize $\hat{c}_i^p$, as shown in formula (4):

$$
\hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_{j} \exp(c_j^p)}
$$

(4)
3. Improved SSD network

3.1. Depthwise separable convolution

Redundant parameters cause slow operation because there are many parameters in the VGG16 network. Therefore, we used MobileNet to replace VGG-16 basic network. MobileNet [12] is a lightweight model, which was proposed by Howard in 2017. The network model consists of 13 convolutional layers, average pooling layer, fully connected layer, and input and output layers. Its core is a depthwise separable convolution, which can solve traditional convolution into a 3×3 depthwise convolution and a 1×1 pointwise convolution. Each convolution layer is followed by a BN layer and a ReLU layer. This reduces model parameters and the amount of calculation. The depthwise structure is shown in Fig.2.

![Fig.2 Depthwise convolution structure](image)

MobileNet network implementation is shown in Fig.3. The left picture is the standard convolution, and the right one is the depthwise separable convolution.

![Fig.3 MobileNet network implementation](image)

3.2. MobileNet-SSD

Using MobileNet network to replace VGG-16 basic network is called MobileNet-SSD, and its basic structure is shown in Fig.4. The MobileNet reduces the final average pooling layer and the fully connected layer. It only retains the first 13 layers of separable convolution as the basic network.

![Fig.4 MobileNet-SSD structure](image)

It is assumed that the input feature map size is $D_F \times D_F \times M$. Where $D_F$ is the width and height of the feature graph. For standard convolution $D_k \times D_k$, the amount of computation is $D_F \times D_F \times M \times N \times D_k \times D_k$. The computation of depthwise convolution is $D_F \times D_F \times M \times D_k \times D_k$. The computation of pointwise convolution is $D_F \times D_F \times M \times N$. So the sum of depth convolution and standard convolution is formula (5).
When $N$ is large, depthwise separable convolution can reduce the amount of calculation by about 9 times compared with standard convolution. Therefore, choosing MobileNet can reduce the running time of the network significantly. The detailed network parameters are shown in Table 1.

### Table 1 MobileNet network parameters

| Network layer / Step size | Input size | Convolution kernel size |
|---------------------------|------------|-------------------------|
| Conv/s2                   | 300*300*3  | 3*3*3*32                |
| Conv dw/s1                | 150*150*32 | 3*3*3*32 dw             |
| Conv/s1                   | 150*150*32 | 1*1*32*64               |
| Conv dw/s2                | 150*150*64 | 3*3*64 dw               |
| Conv/s1                   | 75*75*64   | 1*1*64*128              |
| Conv dw/s1                | 75*75*128  | 3*3*128 dw              |
| Conv/s1                   | 75*75*128  | 1*1*128*128             |
| Conv dw/s2                | 38*38*128  | 1*1*128*256             |
| Conv/s1                   | 38*38*256  | 3*3*256 dw              |
| Conv dw/s1                | 38*38*256  | 1*1*256*256             |
| Conv/s1                   | 38*38*256  | 1*1*256*512             |
| Conv dw/s2                | 19*19*256  | 1*1*256*1024            |
| Conv/s1                   | 19*19*512  | 3*3*512 dw              |
| Conv dw/s2                | 19*19*512  | 3*3*512 dw              |
| Conv/s1                   | 10*10*512  | 1*1*512*1024            |
| Conv dw/s2                | 10*10*512  | 3*3*1024 dw             |
| Conv/s1                   | 10*10*1024 | 1*1*1024*1024           |

### 4. Experiment and result analysis

#### 4.1. Data set and test platform

In order to detect the hot spots of PV array, the corresponding data set should be collected first. Our experimental team collected a total of 2,000 photos of PV arrays, and 1500 of them had hot spots. We also performed data enhancement and expansion on the original data, including flipping, cropping, scale transformation, denoising, and so on. Finally, nearly 6,000 databases were obtained, and 200 were used as the test set. The image comparison after data enhancement is shown in Fig.5.
Label the new data set with labelimg. The CPU is Core i7 processor, the GPU is NVIDIA 1080ti, and the deep learning framework is tensorflow. The model training was completed through the above-mentioned test platform.

4.2. Model training and results
In the training of the MobileNet-SSD network, it is found that the training is difficult, and the network convergence speed is slow or even does not converge. Therefore, this paper adopts the transfer learning method. The common data set Imagenet is used as the pre-training model to share the potential common features. Then use the self-made target data set to train the improved SSD target detection model, then use the test model to verify the speed and accuracy. The detection results of big, medium and small targets for PV hot spot detection are shown in Fig.6.

![Fig.6 Photovoltaic hot spot detection results](image)

It can be seen from Fig.6 that the improved SSD algorithm in this paper has a confidence of 100% for big targets, which has achieved very good results. It also has a confidence of 85% for the middle target and 54% for a small target.

4.3. Model comparison and analysis
In order to verify the effectiveness of the improved algorithm, we compared the classic YOLO algorithm and SSD algorithm. Similarly, perform model pre-training on the basic network of each target detection model. The training and testing of the experimental model adopts the target data set of the laboratory. Under the same training environment and training method, the final test comparison results are shown in Table 2.

| Algorithm | FPS | mPA(%) | Missed detection rate |
|-----------|-----|--------|-----------------------|
| YOLO      | 45  | 82.8   | yes                   |
| SSD       | 46  | 89.2   | no                    |
| Improved SSD | 27  | 87.8   | no                    |

From Table 2, we can see that the MobileNet-SSD simplifies the network structure of the SSD and FPS is been significantly improved from 46FPS to 27FPS, compared with the classic YOLO and SSD. In terms of confidence, the algorithm used in this paper has a big target confidence of 100%. Compared with the YOLO algorithm, the confidence is improved by 5.0%, which is roughly the same as the SSD algorithm. At the same time, in view of the current situation that the current SSD algorithm and the classic YOLO algorithm are difficult to detect small targets, the design has a low missed detection rate and high robustness.

5. Conclusion
In this work, we focused on the deep learning framework in PV hot spot detection, and used MobileNet to optimize the SSD network structure. Experimental results show that the designed PV array hot spot detection system has a good degree of confidence, high detection speed, slow missed
detection rate, and good robustness, compared with the traditional YOLO and SSD algorithms. This has a certain value for the operation and maintenance of smart PV systems. However, this design does not consider the hot spot classification detection. Therefore, the intelligent size detection of the PV hot spot needs to be improved in the subsequent development.

Acknowledgments
This work was financially supported by the Special Project for Transformation and Guidance of Scientific and Technological Achievements of Shanxi Province(201804D131038), and Key R & D Projects of Shanxi Province(201903D421032).

References
[1] Madeti S R, Singh S N. Monitoring system for photovoltaic plants: A review[J]. Renewable and Sustainable Energy Reviews, 2017, 67:1180-1207.
[2] Singh, G. K. Solar power generation by PV (photovoltaic) technology: A review[J]. Energy, 2013, 53:1-13.
[3] Huerta H. L, Pliego M, Alberto, et al. Photovoltaic plant condition monitoring using thermal images analysis by convolutional neural network-based structure[J]. Renewable Energy, 2020, 153:334-348.
[4] M. Alajmi, K. Awedat, M. S. Aldeen and S. Alwagdani. IR Thermal Image Analysis: An Efficient Algorithm for Accurate Hot-Spot Fault Detection and Localization in Solar Photovoltaic Systems[C]// 2019 IEEE International Conference on Electro Information Technology (EIT). Brookings: IEEE, 2019:162-168.
[5] Oliveira A K V D, Aghaei M, Madukanya U E, et al. Aerial Infrared Thermography of a Utility-Scale PV Plant After a Meteorological Tsunami in Brazil[C]// 7th World Conference on Photovoltaic Energy Conversion (WCPEC-7). Waikoloa Village: IEEE, 2018:684-689.
[6] Jiang Lin, Su Jianhui, Shi Yong, Lai Jidong, et al. Thermal spot detection method of photovoltaic array based on infrared thermal image processing[J]. Acta solar energy, 2020, 41(08):180-184(in Chinese).
[7] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// IEEE Conference on Computer Vision & Pattern Recognition. Las Vegas: IEEE Computer Society, 2016:770-778.
[8] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// Computer Vision & Pattern Recognition. Las Vegas: IEEE, 2016:779-788.
[9] J. Redmon and A. Farhadi. YOLO9000: Better, Faster, Stronger[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017:6517-6525.
[10] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision. Springer, Cham: Computer Science, 2016:21-37.
[11] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. IEEE Computer Society, 2016:770-778.
[12] Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications (2017-04-06). https://arxiv.org/pdf/1704.04861.pdf.