Abnormal target tracking and localization algorithm for UAV PV inspection scenarios

XiaoPing Xie*, Xiangui Wei, Xingyu Wang, Xincheng Guo, Ju Li, Zhifeng Cheng
Longyangxia Solar Power, Co., Ltd.of Huanghe Hydropower, Qinhai, China

*Corresponding author e-mail: 316246645@qq.com

Abstract: In order to adapt to the development of UAV inspection technology and achieve accurate positioning of abnormal targets of photovoltaic panels, this paper applies deep learning to abnormal target location, and proposes abnormal target tracking and localization algorithm for UAV PV inspection scenarios. The advantage of deep learning lies in the ability to learn the general expression of the target. In order to obtain an efficient deep network, this paper uses a large amount of data to train the deep network, so that the features extracted by the network can achieve accurate representation of the tracking target. By testing in the abnormal target location test database, it is proved that the proposed algorithm can accurately locate the abnormal target.

1. Introduction

With the development of the economy, the demand for electricity is also becoming more and more intense. As the most important application method of clean energy, photovoltaic power plants have greatly alleviated the problem of electric energy. However, due to the wide distribution of photovoltaic power stations, many types of applications, low efficiency of operation modes, high error rate, long inspection cycle, and even unable to achieve effective operation and maintenance, the healthy development of photovoltaic power plants has been severely restricted [1].

The emergence and rapid development of drones has become one of the effective means to solve the operation and maintenance of photovoltaic power plants. The intelligent drone [2] has the characteristics of high maneuverability, freedom of movement without terrain restrictions, and a variety of portable mission equipment. Used in power system inspection, it can feedback the results in time, automatically identify defects, and effectively predict the failure phenomenon of the power station, which can significantly improve the inspection frequency and inspection efficiency and accuracy, and avoid the inspection process.

![Fig.1 Photovoltaic panel](image_url)
In order to better realize the accurate identification and positioning of abnormal targets in photovoltaic panels shown in Fig.1, this paper proposes an abnormal target tracking and localization algorithm for UAV PV inspection scenarios. In this paper, deep learning is applied to anomaly recognition. The network parameters are trained offline by a large amount of data to learn the general characteristics of abnormal targets. In the process of identification, the depth discriminant features of abnormal targets are extracted to perform real-time target location.

2. Related work

A. UAV inspection

Many domestic enterprises have begun to intervene in the PV inspection market [3], from emerging UAV companies to traditional PV inspection companies and power companies. For UAV enterprises, the advantage is that they can quickly integrate UAV platform, image transmission and PV camera mounting. The professionalism of UAVs is strong. The disadvantages are that they are mostly oriented to multiple industries and do not penetrate into the PV segment. The solution is usually the initial stage, that is, the drone + infrared spot, and there is no room for improvement in efficiency. For traditional photovoltaic testing enterprises and power companies, the advantage is that they have stable market channels, and can directly introduce products and services for drone inspections to customers. The disadvantage is that the new industry of drones is relatively shallow.

In foreign countries, there are also drone infrared inspections and photovoltaics, and the entire market segment is still very new. Some domestic and foreign companies represented by Workswell have proposed the concept of identification + positioning, which can greatly reduce the labor cost, and the implementation methods are different.

B. Neural network and deep learning

The birth of neural networks comes from the study of the human brain. It is well known that there are more than 100 billion neurons in the adult brain, and they bear the signal transmission inside the brain [4]. Biologists revealed the structure of neurons in 1904, and then McCulloch and Pitts referenced the structure of neurons to design the famous MP model [5], artificial neural network was born. After decades of continuous development, the technology of artificial neural networks has matured and the network structure has been diversified, but it has never changed its roots. Fig.2 shows the basic structure of neural networks.

![Fig.2 The basic structure of neural networks.](image_url)

3. Abnormal target location model

A. Convolution feature extraction

Convolution feature extraction is an effective method for processing large images. This method is
based on the inherent characteristics of the statistical properties of any part of the natural image and other parts. The feature learning is performed by CNN[6], since CNN has weights. Sharing the characteristics of the network structure, therefore, compared to the whole image as training data, convolution feature extraction can greatly reduce the scale of the neural network, and at the same time obtain a certain translation, scale and rotation invariance.

The basic process of convolution feature extraction is to first acquire a plurality of smaller partial images by taking the original image, and then use these partial images as training data to form a training set, and perform training operations such as whitening and other processing operations. The neural network uses the back propagation error algorithm to train the network parameters. Finally, the learned network is used as a filter to convolution filter the whole original image to obtain the corresponding convolution characteristics.

B. Network model

In order to obtain the depth discriminant feature, this paper applies deep learning to the algorithm. The deep network used in this paper is the AlexNet network. The AlexNet network [7] is a type of convolutional neural network model. In 2012 Alex Krizhevsky took the AlexNet model he designed and participated in the ImageNet competition and won the championship [8]. Since then, AlexNet has made a name for himself. It also proves the effectiveness of convolutional neural networks in complex models, and the effectiveness of big data + complex network + GPU training in solving problems, which has a milestone significance. Fig.3 is a structural diagram of the target positioning network model.

C. Loss function construction

The strength of the deep network is that it can express the characteristics of the target through a large number of data, and then promote it, extract features in the non-training library, and accurately determine the target. The feature representation ability of the deep network is closely related to its loss function. The loss function represents the task of the deep network. The feature expressions of the depth features extracted by the depth network trained by different loss functions are different.

In this paper, the UAV PV inspection is used as the background, and the deep network conforming to the PV inspection is obtained for training. The logistic function is used as the loss function of this paper.

\[
I(y, x, z) = \log(1 + \exp(-yu(x, z)))
\]

(1)

In the middle, \(x\) is the target template, and \(y\) is the target to be tracked, \(z\) is the search area, and its size is several times the size of the target template, \(u()\) is a function for calculating the degree of
similarity between the target template and the search area. This paper selects n sets of image pairs from the anomalous images and brings them into the deep network to train the network parameters.

$$\text{Loss} = \sum_{i=1}^{n} l(y_i, x_i, z_i)$$

(2)

The training of the network parameters is realized by the loss function to meet the requirements of the task. Through the loss function training, the network can extract the discriminative features of the target.

**D. Optimization solution**

Under the premise of a given tracking target, the abnormal target positioning task realizes continuous and effective tracking of the target in subsequent frames. How to train the network and make the features extracted by the network accurately express the characteristics of the tracking target is the problem that needs to be solved in this paper. According to the characteristics of UAV inspection, this paper establishes the loss function of abnormal target location, and achieves the purpose of training network by deriving the loss function gradient. The input of this paper is a set of sample pairs: the target template $x_i$ and the search area $z_i$. By convolving the target template with the search area, the scores of all candidate samples in the search area are obtained.

$$u(x_i, y_i) = \phi(x_i) * \phi(z_i)$$

(3)

In the middle, $\phi(.)$ is the deep network used in this paper. The depth discriminative features of the target template and the search area are extracted through the deep network, and convolved to obtain the degree of similarity between the target template and the search area. The similarity score is compared with the label to optimize the network parameters.

$$\arg \min_{\phi} \sum_{i=1}^{n} \log(1 + \exp(-y_i u(x_i, z_i)))$$

(4)

By minimizing the loss function, the similarity score of the target template and the search area is consistent with the label, thereby achieving the goal of meeting the task.

4. **Experiment**

**A. Algorithm verification**

In order to better obtain the deep discriminative features that can express the target, before the test, the network is trained with the abnormal target video. In order to prove the effectiveness of pre-training, two sets of experiments were carried out on pre-training respectively. Among them, OURS_YP is the result obtained by the pre-training algorithm, and OURS_NP is the result obtained by the algorithm without pre-training. The result is as follows:

As shown in the Fig.4, the pre-training algorithm has an accuracy of 0.728 and a success rate of 0.542, both of which are unpre-trained. The anomaly target location of this paper is different from the mission goal of the usual target tracking. The pre-training process is to enable the network to locate this special task according to the abnormal target, and to correct the network parameters, so that the features extracted by the network can express abnormal targets more. Through experiments, it is proved that pre-training is effective.
B. Overall performance assessment

In order to prove the advanced nature of the algorithm, this paper compares the algorithm with the latest target tracking algorithm. The comparison algorithms are HCF\[9\], DSST\[10\], SINT\[11\], KCF\[12\] and SAMF\[13\]. In this paper, the compared tracking algorithm is tested in the abnormal target test database, and the test results are obtained. Then the test results of this paper are compared with the test results of the comparison algorithm. The results are shown below.

It can be seen from the Fig. 5 that compared with the comparison algorithm, the accuracy of the proposed algorithm is 0.728, and the success rate is 0.542, which is higher than the HCF algorithm in the second place, and the best tracking and positioning performance is obtained. Through experiments, the algorithm is effective and can accurately track abnormal targets.

C. Algorithm speed comparison

For the task specificity of this paper, the speed of the algorithm plays an important role in the completion of the task. In this paper, the real-time performance of the algorithm is used as the basis for discriminating the performance of the algorithm. Compared with the current tracking algorithm, the comparison algorithms are HCF, DSST, SINT, KCF and SAMF. The algorithm speed table is shown in Table 1:

![Fig.5 Overall performance comparison chart](image)

| Algorithm  | OURS | HCF | KCF | DSST | SINT | SAMF |
|------------|------|-----|-----|------|------|------|
| FPS        | 30   | 11  | 245 | 24   | 1    | 20   |

It can be seen from the table that the speed of the algorithm is 30fps, which meets the requirements of real-time. Although the algorithm is slower than KCF, it still meets the requirements of the task, and the accuracy of the algorithm is better than KCF. Overall, the performance of the algorithm reaches the top level.

5. Conclusion

In order to cooperate with the target UAV inspection technology, and better enable the UAV to accurately identify and locate the abnormal targets in the PV panel, this paper proposes an abnormal target tracking and localization algorithm for UAV PV inspection scenarios. In this paper, the deep discriminative features of the tracking target are extracted through the deep network to achieve accurate positioning of the abnormal target. At the same time, the experiment is carried out in the abnormal target test database, which proves the effectiveness of the proposed algorithm.

References

[1] Bo C , Xiao-Dong Z , Ling-Zhi Z , et al. Strategy for reactive control in low voltage ride through of photovoltaic power station[J]. Power System Protection and Control, 2012.
[2] Yavuz D , Akbiyik H , Bostanci E . Intelligent drone navigation for search and rescue operations[C]. 2016 24th Signal Processing and Communication Application Conference (SIU). IEEE, 2016.
[3] Wu Y , Zhao G, Hu J, et al. Overhead Transmission Line Parameter Reconstruction for UAV Inspection Based on Tunneling Magnetoresistive Sensors and Inverse Models[J]. IEEE Transactions on Power Delivery, 2019, PP(99):1-1.
[4] Singh G, Gehr T, Püschel M, et al. An abstract domain for certifying neural networks[J]. Proceedings of the ACM on Programming Languages, 2019, 3(POPL):1-30.
[5] Mcculloch W S, Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity[J]. Bulletin of Mathematical Biology, 1943, 52(1–2):99-115.

[6] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6):1137-1149.

[7] Gu S, Ding L, Yang Y, et al. A new deep learning method based on AlexNet model and SSD model for tennis ball recognition[C]. 2017 IEEE 10th International Workshop on Computational Intelligence and Applications (IWCLA). IEEE, 2017.

[8] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural Networks[C]. NIPS. Curran Associates Inc. 2012.

[9] Ma C, Huang J B, Yang X, et al. Hierarchical convolutional features for visual tracking[C]. Proceedings of the IEEE international conference on computer vision. 2015: 3074-3082.

[10] Danelljan M, Häger G, Khan F, et al. Accurate scale estimation for robust visual tracking[C]. British Machine Vision Conference, Nottingham, September 1-5, 2014. BMVA Press, 2014.

[11] Tao R, Gavves E, Smeulders A W M. Siamese instance search for tracking[C]. Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1420-1429.

[12] Henriques J F, Caseiro R, Martins P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE transactions on pattern analysis and machine intelligence, 2014, 37(3): 583-596.

[13] Li Y, Zhu J. A scale adaptive kernel correlation filter tracker with feature integration[C]. European conference on computer vision. Springer, Cham, 2014: 254-265.