SCIENCE CHINA
Information Sciences

• NEWS & VIEWS •

April 2023, Vol. 66 140307:1–140307:2
https://doi.org/10.1007/s11432-022-3663-1

Special Topic: Artificial Intelligence Innovation in Remote Sensing

# AIR-PV: a benchmark dataset for photovoltaic panel extraction in optical remote sensing imagery

Zhiyuan YAN[1,2,3], Peijin WANG[1,2,3], Feng XU[4], Xian SUN[1,2,3] & Wenhui DIAO[1,2,3*]

[1]*Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing* 100190*, China;*
[2]*School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing* 100049*, China;*
[3]*Key Laboratory of Network Information System Technology (NIST), Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing* 100190*, China;*
[4]*Key Laboratory for Information Science of Electromagnetic Waves (MoE), Fudan University, Shanghai* 200433*, China*

Received 31 August 2022/Revised 20 November 2022/Accepted 19 December 2022/Published online 17 March 2023

**Citation** Yan Z Y, Wang P J, Xu F, et al. AIR-PV: a benchmark dataset for photovoltaic panel extraction in optical remote sensing imagery. Sci China Inf Sci, 2023, 66(4): 140307, https://doi.org/10.1007/s11432-022-3663-1

The energy crisis has been a hot topic of global concern. Photovoltaic devices, a typical new energy source, have progressed rapidly and become among the main sources of power generation in the world [1]. Most photovoltaic fields are often constructed in large areas, making it difficult to monitor photovoltaic panel situations. Combining remote sensing (RS) and deep learning [2], using algorithms to automatically monitor the status of photovoltaic power plants and their distribution from satellite images provides an effective way for energy monitoring and assessment and can effectively aid policymakers.

However, extracting photovoltaic panel regions from RS images is a challenging task. On the one hand, the RS image background is complex. There exist some objects that are easily confused with photovoltaic panels, such as dark buildings and roads. On the other hand, photovoltaic panels are often sparsely distributed with data imbalance problems, making it hard to achieve accurate extraction.

Several recent studies on photovoltaic panel extraction have emerged in the RS field. Zhao et al. [3] proposed a method with a small dataset for photovoltaic panel detection based on aerial drone data. Song et al. [4] presented a dataset for photovoltaic land detection in RS scenes with deep learning and promoted automated photovoltaic panel extraction in the RS field. However, some limitations still remain. (1) The spatial distribution of the dataset is concentrated in a local area, leading to small data volume and simple backgrounds. While deep learning models, which are data-driven, usually require a large amount of data with refined ground truth to obtain a robust performance. (2) These datasets are not publicly available, which cannot further facilitate more research in this field.

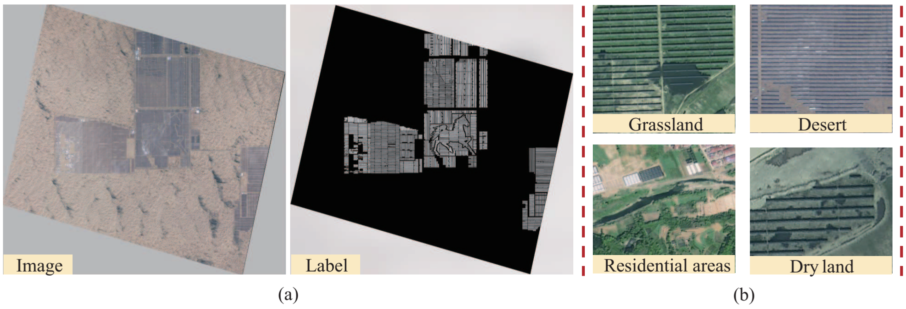
Herein, a large-scale benchmark dataset for photovoltaic panel extraction in RS imagery, named AIR-PV, is proposed. Figure 1 shows some examples. The main characteristics of our dataset can be summarized as follows. (1) The dataset is large-scale with wide distribution. The photovoltaic regions in AIR-PV are spread across five provinces of western China to cover a wide range of geographical styles and background diversity. In addition, it covers more than 3 million square kilometers with more than 300000 photovoltaic panels in the dataset. (2) The dataset is publicly available. AIR-PV is one of the earliest publicly available datasets for photovoltaic panel extraction in RS imagery. It aims to provide a standard data foundation for applying advanced deep learning technology to photovoltaic panel extraction in RS, thereby promoting various social applications related to photovoltaic power.

*Data collection and preprocessing.* Because of the vigorous development of the photovoltaic industry in Western China, data from five representative provinces, namely, Inner Mongolia, Qinghai, Gansu, Shanxi, and Yunnan, were collected. The world's largest "horse" photovoltaic power station in Inner Mongolia was also included in AIR-PV, as shown in Figure 1(a). The images were from the Gaofen-2 satellite with 0.8 m resolution. As shown in Figure 1(b), photovoltaic areas with backgrounds of desert, grassland, residential area, and dry land were collected. The size of the original RS images was more than 15000 × 15000 pixels. Deep models cannot be directly trained on such a large size. Therefore, the images were cropped into slices with 512 × 512 pixels. In this way, 5000 samples were obtained.

*Image annotation.* We treat the photovoltaic panel extraction as a semantic segmentation task. The annotation follows the standard format of MSCOCO [5] segmentation, which is a PNG file. There are two categories including photovoltaic panels marked in white and the others marked in black with pixel level. Annotation quality was guaranteed through a strict annotation process.

*Baseline experiments and analysis.* Baseline methods SegNet [6], PSPNet [7], and Deeplab-v3+ [8] were adopted to evaluate the performance of AIR-PV. 70% of the dataset

* Corresponding author (email: diaowh@aircas.ac.cn)




| Methods | mIOU (%) |
|---|---|
| SegNet | 80.58 |
| PSPNet | 85.21 |
| Deeplab-v3+ | 86.74 |

(c)

**Figure 1** (Color online) Examples of images and annotations of AIR-PV. (a) Visualization of a large scene in the AIR-PV dataset. The "horse" photovoltaic power station in Inner Mongolia. (b) Photovoltaic under the background of different ground objects. (c) Results of baseline models.

was randomly selected for training and 30% for testing. mIOU was used to evaluate the model's accuracy. Figure 1(c) exhibits the results. Deeplab-v3+ achieves a higher accuracy of 86.74%, which outperforms those of PSPNet (1.53%) and SegNet (6.16%). This may be because Deeplab-v3+ incorporates the shallow features that contain more localization and structure information of photovoltaic panels in the model. Moreover, it can capture more contextual information through ASPP to assist photovoltaic panel extraction under different backgrounds. Generally, there is still room for improvement in accuracy, which reveals the challenge of our dataset.

*Conclusion*. In this work, a large-scale benchmark dataset for photovoltaic panel extraction in RS imagery was proposed. Experiments were conducted for performance evaluation of baseline methods. The results demonstrated the challenges of the dataset. Future work is needed to expand the data collection area and increase the regional richness of the dataset. This work reveals that AIR-PV will help in promoting technological development in this field, thereby promoting global energy analysis as well as various potential social applications.

*Access dataset*. The datasets and more detailed introduction can be obtained from the website https://github.com/AICyberTeam.

**Acknowledgements** This work was supported by National Natural Science Foundation of China (Grant Nos. 62171436, 61725105).

**References**

1 Zhao J, Chen L, Wang Y N, et al. A review of system modeling, assessment and operational optimization for integrated energy systems. Sci China Inf Sci, 2021, 64: 191201
2 Xu F, Hu C, Li J, et al. Special focus on deep learning in remote sensing image processing. Sci China Inf Sci, 2020, 63: 140300
3 Zhao L L, Wang Y, Liu J. Detection and analysis of photovoltaic panels based on UAV and HSV space. Infrared Technology, 2020, 42: 978–982
4 Song Y C, Li Y C, Geng Z Y, et al. Application of deep learning method in remote sensing detection of photovoltaic land (in Chinese). Science Surveying Mapping, 2020, 45: 84–92
5 Lin T, Maire M, Belongie S, et al. Microsoft COCO: common objects in context. In: Proceedings of European Conference on Computer Vision, 2014. 740–755
6 Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell, 2017, 39: 2481–2495
7 Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2017. 17355193
8 Chen L, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of European Conference on Computer Vision, 2018. 833–851