

This research focused on validating deep learning as a tool to automatically extract photovoltaic panels from satellite and aerial imagery with various spatial resolutions. Using RGB bands from images, the study received a high accuracy in the classification and segmentation of PV panels, and the transferability of the models trained with different resolution samples was also discussed. Generally, the proposed approach was appropriate for the aim of this study. I have some concerns and suggestions.

Response: Thank you for your encouraging comments.

Comments:

1 – According to the classification system introduced in Table 1, the PV dataset introduced in this study may including different categories of PVs. However, in the result part, all the categories were classified as a whole. The authors may want to show the influence of different resolutions on classifying different categories of PVs, but the “classification system” used here may confuse. My suggestion is to change the term “classification system” in the title of Table 1.

Response: Thank you for your suggestion. In order to support more applications and give full play to the data value, we divide all samples into different sub-categories according to their background land use or roof type. In the result part, we focus on validating the impact of different resolutions on PV classifications; hence, all sub-categories are classified as a whole. To avoid confusion, the term “classification system” is replaced by “organizational structure” in the revised manuscript, that is, the revised title of Table 1 is “Organizational structure of our PV dataset.”

2 – The names of segmentation networks in the manuscript should be checked. For example, “DeepLab v3+” was taken as “DeepLab v3” in Line. 189, Page. 11.

Response: Thank you very much. We have carefully checked and corrected the names of the three segmentation networks in the manuscript. “DeepLab v3” should be “DeepLab v3+” and the revised sentence is “On average, DeepLab v3+ achieved an IoU of 0.873 for flat concrete PVs and 0.927 for steel tile PVs.”

3 – This study gave suggestions on the selection of image resolutions for the classification of different PVs. The differences in the classification results of images with different resolutions may be related to the size of the target features and input samples, because the semantic segmentation networks are generally sensitive to the size, shape, and receptive fields. It may be interesting to give a quantified result on selecting image resolutions and input sample sizes for target features with different sizes in the future study.

Response: Thank you for your suggestions. In this study, we recommend to use PV08 for concentrated PV, PV03 for distributed ground PV, and PV01 for distributed rooftop PV so as to achieve the best segmentation results. However, the quantitative result on selecting image resolutions and input sample sizes for PVs with different sizes might require samples from more satellite and aerial images with different resolutions. Therefore, this interesting suggestion is only discussed in Section Conclusion of the revised manuscript. The related part is “Besides, this dataset may contribute to a diversity of other research and applications related to PV. For example, the segmentation networks are generally sensitive to the observational size and shape in the receptive field; hence, it is valuable to quantitatively explore the general guidelines on selecting image resolutions and input sample sizes for PVs with different sizes.”

4 – Authors used images with different resolutions to extract PVs in the study area. However, it may be interesting to see the classification results when combining multiscale features from images with different resolutions. The fusion of multiscale features can be further discussed in the future.

Response: Thank you for your suggestions. This paper presents two simplest applications based on our PV dataset. Now that multi-resolution dataset has been established, more complex cases are feasible. We hope that in the future our dataset will support more valuable and interesting research, for example, investigating whether a network can be established to fuse multiscale features from images with different resolutions to achieve synchronous identification or segmentation of multi-scale PVs. We have added this point into the Conclusion in the manuscript as “Whether a network can be established to combine images with different resolutions to achieve synchronous identification or segmentation of multi-scale PVs is also of great interest.”