

Earth Syst. Sci. Data Discuss., referee comment RC3
<https://doi.org/10.5194/essd-2021-270-RC3>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.



Comment on **essd-2021-270**

Anonymous Referee #3

Referee comment on "Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery" by Hou Jiang et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-270-RC3>, 2021

This paper introduces a multi-resolution PV dataset that composes of various samples collected from satellite and aerial images. Such a dataset is of great interest to users because it helps to develop deep learning algorithms for automatic PV information extraction. The paper is well organized and the dataset is described in a clear fashion. Therefore, I would like to suggest for publication in the Earth System Science Data after a minor revision.

comments:

- For the segmentation experiments (Section 4.1), the samples are divided into 80% training set and 20% testing set. Is the division performed randomly for PV08/03/01 or separately for each subcategory? The authors should make it clear because the two approaches lead to different model performance and segmentation results.
- For the cross application (Section 4.2), I wonder whether authors divide training samples and testing samples in the same way as Section 4.1? The authors state that "fine-tuning means that the model was first pre-trained on PV03 (PV08) samples, then fine-tuned (fine-tuning process lasted 10 epochs) using PV08 (PV03) samples, and finally applied to PV08 (PV03) samples." Are the samples used for fine tuning the same as those used for direct training? If yes, how will the model perform if only using a small portion samples for fine-tuning.