This paper presents a dataset as a benchmark for developing and evaluating methods to extract PV panels from satellite images. Such a dataset is scarce but timely for the relevant community and will help researchers reduce the time and cost of collecting high-quality samples and will enrich other work related to solar PV energy. Overall, this manuscript is technically sound and well-crafted.

I would like to suggest a revision before it can be considered for publication in the Earth System Science Data.

Major concerns:

1. The dataset was collected in Jiangsu Province. Considering that the changes in geographic context may affect the model's performance, how can you guarantee that the model trained with this dataset can be generalized to other regions?

2. The dataset is composed of Gaofen-2 and Beijing-2 imagery, which include a near-infrared band in addition to RGB bands. Does this mean that the model trained with this dataset can only be applied to the satellite imagery with a specific combination of spectrums? Are these two imagery representative of the current common satellite imagery?