2024 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm) | 979-8-3503-1855-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/SmartGridComm60555.2024.10738071

Mapping Solar Potential: Comprehensive PV System Detection with Aerial and Satellite Images Using SolarSAM

1st Jessy Matar
ASCI
fortiss GmbH
Munich, Germany
matar@fortiss.org

2nd Michael Albrecht
ASCI
fortiss GmbH
Munich, Germany
albrecht@fortiss.org

3rd Aleksandre Kandelaki
Technical University of Munich
Munich, Germany
kandelakialeksandre@gmail.com

4th Florian Kotthoff
Division of Energy
OFFIS
Oldenburg, Germany
florian.kotthoff@offis.de

5th Markus Duchon
ASCI
fortiss GmbH
Munich, Germany
duchon@fortiss.org

***Abstract*—Automatic recognition of photovoltaic (PV) systems through remote sensing is critical for energy and infrastructure planning. This study explores the efficacy of deep learning in detecting PV systems using remote sensing. We introduce the adaptation of the Segment Anything Model (SAM) to this task, marking the first application of SAM for the detection and delineation of PV installations. We train the model on high-resolution images and then test its performance on both high-resolution aerial images and lower-resolution satellite images. Our results demonstrate high detection accuracy and precision especially for aerial imagery. We underscore the potential of advanced deep learning techniques in detecting and monitoring PV installations, facilitating more effective planning and deployment of renewable energy resources. Our findings suggest that the application of SAM in this domain could lead to significant advancements in energy infrastructure development, offering a promising tool for optimizing the integration and management of solar energy systems on a global scale.**

***Index Terms*—Photovoltaic, Deep Learning, Remote Sensing, Solar Energy**

I. Introduction

Automated detection of photovoltaic (PV) installations in urban and open areas is a critical task with implications for energy planning and infrastructure development. Remote sensing imagery offers a promising avenue for this purpose, especially for effective integration of distributed renewable solar resources into existing power grids (Malof et al. [2019], Kruitwagen et al. [2021]). We seek to address questions surrounding the influence of resolution on deep learning models' efficacy, investigating how the performance is affected when trained on high-resolution and tested on lower resolution images. From a deep learning and computer vision perspective, exploring the impact of image resolution on model performance is particularly intriguing. Computer vision algorithms have become instrumental in remote sensing applications tasks (Illarionova et al. [2023]). Remote sensing images offer potential for identifying solar installations (Bradbury et al. [2016], Aliabad et al. [2022]), aiding in the detection of newly constructed PV systems, estimating and planning of rooftop PV installation, and power generation forecasting. This enhanced monitoring capability facilitates informed decision-making and allows the creation of databases of renewable energy systems at a large scale (Mayer et al. [2020]). Despite the promise of remote sensing imagery, challenges persist, including limited availability of annotated datasets and the specificity of PV installation shapes. Moreover, the evolving nature of PV installations necessitates continual updates, rendering previous annotations obsolete. Semantic segmentation, involving the assignment of semantic labels to each pixel in an image, offers a potential solution but requires robust training data. Various deep learning (DL) models have been developed to detect PVs (Yu et al. [2018a], Hou et al. [2020], Ortiz et al. [2022]), however it remains challenging to segment PV systems globally due to the lack of global high resolution coverage and the difference in the PV systems characteristics. Hence, the need for a robust DL model that can advance detection, which would be capable of handling imagery with varying resolutions captured by different cameras. We use an emerging high performing flexible model, SAM (Kirillov et al. [2023]), and adapt it to detect rooftop and solar plants PV installations. Ren et al. [2023] deduced that SAM is sensitive to the resolution of the imagery which can significantly affect its performance. By investigating various image resolution, we aim to test the model's predictions and effectively assess the model's performance.

II. Data Collection

In this section, we present our approach to collecting data for analyzing PV installations, considering both aerial and satellite imagery. Aerial data sources offer high-resolution images, allowing for precise detection of buildings and rooftops. On the other hand, satellite imagery provides broader regular coverage but typically at a lower resolution. The preparation of the dataset plays a crucial role in determining the quality of training and performance.

A. Aerial Imagery From OpenData

OpenData portal by the Bavarian Surveying Administration provides access to free geospatial datasets for diverse applications. The repository provides access to high-quality datasets, ranging from geographical information to environmental data. We obtain aerial data from OpenData, which provides imagery with resolutions as high as 40 cm and covering a wide region in Germany. The masks were manually created for the images using QGIS. It is in geojson format, which is widely used to geo-spatially specify geometries and save them for later use. It provides wide variety of geometries, such as, Polygons, multi-polygons, Points, Lines etc. The masks used for training were created via the Multi-polygon geometry.

Fig. 1. Image and respective ground truth Mask cropped into 256x256 patch.

Figure 1 visualizes an example image from the training dataset with respective mask. With such high resolution, buildings and rooftops with PV installations can be easily distinguished and detected.

B. Satellite Imagery From Sentinel-2

Sentinel-2, is part of the European Union's Copernicus program. It offers freely accessible satellite imagery through online data hubs, capturing Earth's surface with high-resolution optical sensors (Immitzer et al. [2016]). The satellite carries the Multispectral Imager (MSI), which sensors deliver 13 spectral bands ranging from 10 to 60-meter pixel size. Sentinel-2 data enables diverse applications like land cover mapping, agriculture monitoring, and environmental assessment. It offers a global coverage and provides frequent revisits, which empowers users worldwide to gather valuable insights for research and planning. Despite having a resolution of 10 meters, Sentinel-2 can identify PV installations, although its capability is primarily suited for large solar farms. PV roof installations typically have dimensions smaller than 10 meters and are therefore not easily detectable using Sentinel-2 imagery. For an object to be discernible, it needs to be larger than a single pixel, therefore Sentinel-2's resolution is insufficient for accurately delineating residential PV installations. The satellite's resolution is too coarse, restricting its ability to identify smaller PV installations effectively. We collect few sample images from Sentinel-2 to detect solar farms and test our deep learning model capabilities on 10m resolution RGB images.

III. Deep Learning Model: SolarSAM

Deep Learning is a specific form of machine learning that is based on neural networks and is characterized by having multiple layers. To train deep learning models with good generalization abilities, one needs large data sets. This is true for both fine tuning models and training small networks from scratch (Zhu et al. [2017]). Semantic segmentation is an application of DL where the classification of each pixel of an image occurs. In the context of PV installation detection, image segmentation techniques can be employed to delineate the boundaries of individual solar panels within an image, enabling accurate identification and assessment of PV systems.

SAM is a segmentation model developed and maintained by Meta. The primary goal of creating SAM was to build a foundation model for image segmentation task (Kirillov et al. [2023]). Foundation models are characterized with their ability to generalize quite well to more specific tasks with zero-shot or few-shot prompting. Zero-shot prompting allows for prompt-based responses without the need for specific training. In contrast, the few-shot prompting technique enhances the model's performance by employing a limited number of examples. This approach, known as "Prompt Engineering," enables SAM to accurately segment any object in an image when appropriately instructed (Ren et al. [2023]). The prompt could be a mouse click to a specific region of the photo or a text-based prompt when applying zero-shot approach. Alternatively, to enhance performance, it is possible to utilize a pretrained model and further train it using the new dataset. This process causes the model's distribution to align with that of the dataset, resulting in an improved segmentation capability for the objects contained within the dataset.

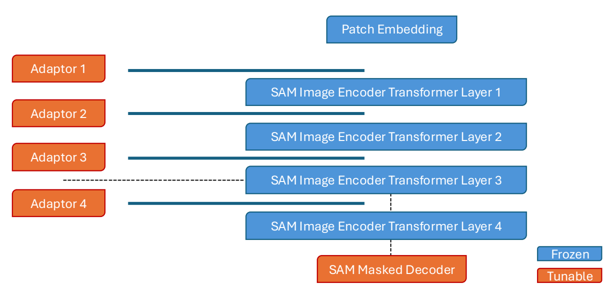

Fig. 2. Ground truth and predictions of rooftop PVs using SolarSAM.

However, the limitations of SAM's coverage become evident when considering the wide array of issues present in computer vision. This is a common issue (or even desired issue) among foundation models, as the training data is unable to capture the entirety of the segmentation problem. SAM model is gaining attraction in performing downstream tasks with great success (Wu et al. [2023]). SAM-Adapter addresses the mentioned flaw by utilizing the concept of visual prompting by using Adaptors (Chen et al. [2023]). In their work, Chen et al. [2023] showed that the SAM-Adapter can significantly elevate

Fig. 3. Ground truth and predictions of rooftop PVs using SolarSAM.

the performance of SAM in challenging tasks as shown in extensive experiments.

The SAM-Adapter algorithm, which utilizes SAM as the backbone, injects customized information into the network through simple yet effective adapters to enhance performance in specific tasks. As a base model adaptable to specific tasks through fine-tuning, we utilize the same simple algorithm as the SAM-Adapter version (includes an image encoder and a mask decoder) and tailor it to our downstream objective and assess its effectiveness. Our application involves employing the SAM-Adapter approach for the segmentation of PV systems which we call SolarSAM.

Figure 2 depicts the architecture of SAM-Adapter and visualizes how and where the visual prompting using Adaptors is applied. The adaptors used, represent a function, which generates a tailored prompt for the PV detection task that we subsequently inject into the network for fine-tuning. The input for the Adaptor is the task-specific knowledge obtained from the new dataset, in our case PV installations. As a result, SolarSAM is fine-tuned for a specific downstream task and is expected to perform better than the general model.

IV. Results: Resolution Detection With SolarSAM

We explore in this section the effectiveness of SolarSAM in accurately delineating PV systems and discuss its potential.

A. Training on High-Resolution Dataset

We trained SolarSAM on a high-resolution dataset to identify and delineate photovoltaic (PV) installations with the aim of evaluating its capacity to handle fine details and varying scales within the images, including both small rooftop PV installations and large solar farms. During the training process, the model was exposed to detailed high-resolution images, allowing it to learn and recognize the intricate features of PV panels, as seen in Figure 3. The total ground truth area for the installed PVs in the shown case corresponds to 4620.30 m^2 while the predicted area is equivalent to 5430.72 m^2 with a relative standard deviation of 18%.

In image segmentation tasks, evaluation metrics such as Intersection over Union (IoU), Precision, and Recall are crucial for assessing model performance (Padilla et al. [2020]). IoU measures the overlap between the predicted segmentation and the ground truth, providing a direct indication of accuracy. Precision reflects the proportion of true positive predictions among all positive predictions made by the model, while Recall indicates the model's ability to identify all relevant instances in the data.

For SolarSAM, we achieved an IoU of 76.06%, a Precision of 88.44%, and a Recall of 83.01% on the aerial images dataset. It's important to note that our training dataset is relatively small compared to those used in previous studies, such as Jiang et al. [2021], Zhuang et al. [2020], and Yu et al. [2018b], which trained models like UNet++ (80cm), DeepSolar (15cm), and CrossNet (30cm) for PV detection

TABLE I
SOLARSAM COMPARED TO STATE-OF-THE-ART; ALL METRICS ARE IN %; (*) USED MULTIPLE RESOLUTIONS FOR TRAINING (10CM, 30CM, 100CM)

Model	Reference	Resolution	IoU (%)	Precision (%)	Recall (%)	Dataset
DeepSolar	Yu et al. [2018b]	15cm	-	93.10	88.5	-
DeepSolar For Germ.	Mayer et al. [2020]	5cm	-	92.66	97.43	-
UNet++	Jiang et al. [2021]	≤80cm (*)	84.40	-	-	-
CrossNet for PV Det.	Zhuang et al. [2020]	30cm	74.26	-	-	135km^2
SolarSAM	Ours	40cm	76.06	88.44	83.01	121.8km^2

Fig. 4. Detection of PV installations using SolarSAM on 40 cm resolution.

on large datasets. Despite this limitation, SolarSAM performs competitively with these established models as seen in Table I. With a larger and more diverse dataset, it has the potential to surpass current state-of-the-art results.

These metrics highlight SolarSAM's strong ability to detect and accurately segment PV installations. Specifically, the model was able to precisely identify individual solar panels and their boundaries, regardless of the installation size or complexity. This is further evidenced by the model's performance across different scales, as shown in Figures 4 and 5.

B. Testing on Low-Resolution Images

To further evaluate SolarSAM's robustness, we tested the high-resolution model on low-resolution images from the Sentinel-2 satellite. This step was crucial to understand how the model performs under different imaging conditions and to assess its adaptability to lower-quality data. Figure 5 shows the model's predictions on these low-resolution images.

As anticipated, the reduction in image quality resulted in a slight decrease in performance. However, despite the

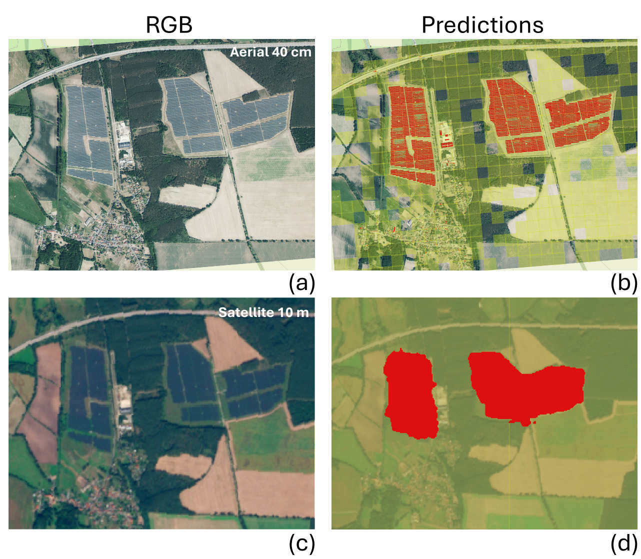

Fig. 5. Same as Figure 4 on 40 cm and 10 m resolution for .

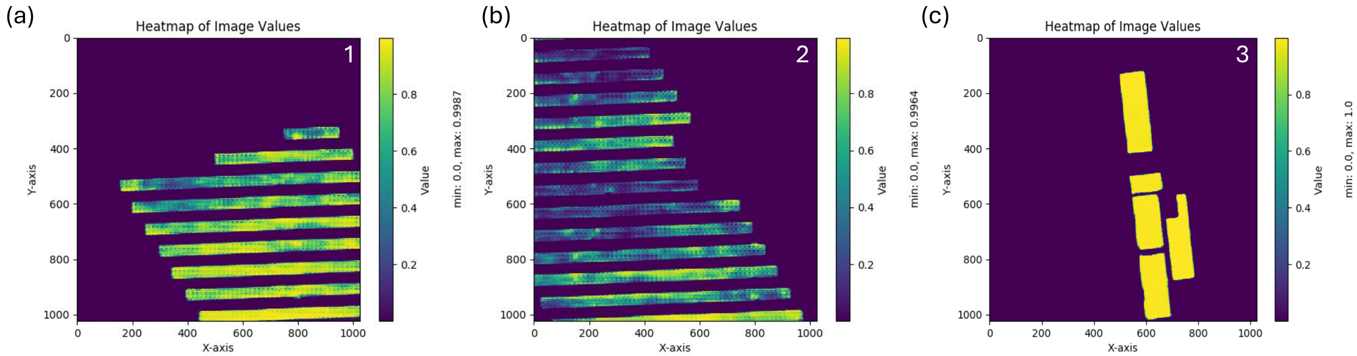

Fig. 6. Confidence predictions using SolarSAM on the high resolution image (graphs annotated 1,2, and 3 to match patches in Fig 4).

diminished detail in the Sentinel-2 images compared to the high-resolution 40 cm imagery, SolarSAM still maintained a commendable level of accuracy. The model successfully detected and delineated the boundaries of solar farms, achieving predictions that, while slightly less precise, remained reliable even with less detailed input.

This capability is further demonstrated by the confidence levels of the predictions, as shown in Figure 6. The model's consistent performance with an IoU that remained above 70% even with lower resolution images underscores SolarSAM's adaptability and effectiveness across diverse imaging conditions. This adaptability is particularly important for large-scale applications, where access to high-resolution data may be limited or costly. The model's proficiency in various conditions suggests that it can contribute significantly to optimizing the integration and management of solar energy systems globally.

V. Discussion

A. Resampling Strategy

In a prior investigation, we wanted to determine the minimum resolution at which the deep learning model can accurately detect and classify photovoltaic systems in images. We generated different resolutions from the same dataset through downsampling then upsampling and then systematically comparing performance across those.

We found that lower resolution bound depends on the use case and context. There might be a generic lower boundary resolution that can work for all the use cases, however finding such boundary demands exploring a huge variety of aerial imagery to ensure that the statements can be broadly applied. The lowest resolution trained and tested on using SAM was 120cm (0.12m), which achieved an IoU score of 54%.

B. Case Study Observations

In our case studies, we conclude that high-resolution imagery is highly effective in identifying both rooftop PV installations and large solar farms, despite SolarSAM being primarily trained on rooftop PVs. This demonstrates the model's potential in handling different types of PV installations. On the other hand, low-resolution imagery proved to be more suited for identifying large-scale solar farms spread over vast areas, as expected.

C. Training Data Bias

Most of the labeled images used for training predominantly cover southern Germany. This geographical bias could potentially affect the model's performance when applied to other regions within the country or to other countries with different environmental and infrastructural backgrounds. Including broader geographical distribution will add background diversity to the model and could improve its performance.

D. Confidence Levels in Predictions

Figure 6 presents the confidence levels of SolarSAM's predictions for three specific patches taken from the broader context shown in Figure 4. These graphs provide a clear visualization of how the model distinguishes between high-confidence (accurate) and low-confidence (less accurate) predictions. High-confidence predictions (closer to value 1) are typically associated with outputs where the model is more certain of the presence and characteristics of photovoltaic (PV) installations, leading to more reliable segmentation and area estimation. Conversely, low-confidence predictions indicate areas where the model is less certain (closer to value 0), which could be due to factors like poor image quality, ambiguous features, or complex roof structures.

E. Future Improvements

Our findings suggest that incorporating a more diverse dataset—including various backgrounds, resolutions, and geographic locations—would significantly enhance the model's Precision and Recall metrics, thereby improving overall performance. However, it is important to note that lower resolution images inherently introduce a higher error rate in area calculation, which impacts the accuracy of PV system capacity estimates.

By addressing these limitations and expanding the diversity of training data, we can further refine SolarSAM's performance, ensuring more accurate and reliable detection and delineation of PV systems across different regions and imaging conditions.

VI. Conclusion

In this work, we explore the SAM-Adapter architecture, to fine-tune SAM for the PV systems detection task. We propose a model for detecting PV installations from aerial and satellite images. We investigate the effectiveness of SolarSAM in accurately delineating PV systems. One significant advantage of SolarSAM is its ability to be used as a monitoring tool of PV systems, taking into consideration different resolutions. It can be used to track the progress of solar PV installations and gathering information of current energy grids relevant to create a stable grid operation.

Focusing on PV installations detection at varying resolutions, we find that high-resolution imagery is optimal for detecting small-scale rooftop installations typically found in residential areas. This yields precise predictions and enables accurate capacity estimates crucial for infrastructure planning. Conversely, while satellite imagery offers lower resolution, it remains abundant and freely accessible, making it a valuable tool for global monitoring despite its limitations in detailed analysis. The trade-off between resolution, cost, and accuracy is a critical consideration, especially when scaling to larger geographical areas.

Moreover, SolarSAM's ability to delineate PV systems with good performance can significantly contribute to the development of predictive models for solar energy generation. Future plans include using SolarSAM to create low-resolution binary masks for PV farms detected from Sentinel-2 images, which will facilitate the use of supervised learning techniques to train models for detecting PV from space. These predictions could provide valuable information, such as precise locations (longitude and latitude), area boundaries, power capacity, energy generation, and other essential data.

SolarSAM stands out not by chasing state-of-the-art complexity or global reach, but by offering a practical, scalable, and cost-effective solution for real-world rooftop photovoltaic (PV) detection. SolarSAM is designed for easy deployment and adaptability to local environments. Its true value lies in bridging the gap between advanced technology and the specific needs of businesses and municipalities, making it a reliable tool that drives actionable insights and accelerates solar energy adoption in urban areas.

Acknowledgment

We gratefully acknowledge financial support through the project executing agency Jülich (PTJ) with funds provided by the Federal Ministry for Economic Affairs and Climate Action (BMWK) due to an enactment of the German Bundestag under Grant No. 03EN3077%. We would also like to acknowledge the European Space Agency (ESA) for providing the Sentinel-2 data used in this study. The Github for the Segmentation using SAM adapter (SolarSAM) code is publicly available (https://github.com/kandelak/PV-Segmentation-SAM-Adapter).

References

F.A. Aliabad, H.R.G. Malamiri, S. Shojaei, A. Sarsangi, C.S.S. Ferreira, and Z. Kalantari. 2022. Investigating the Ability to Identify New Constructions in Urban Areas Using Images from Unmanned Aerial Vehicles, Google Earth, and Sentinel-2. *Remote Sensing* 14, 13 (2022), 3227. https://doi.org/10.3390/rs14133227

Kyle Bradbury, Raghav Saboo, Timothy Johnson, Jordan Malof, Arjun Devarajan, Wuming Zhang, Leslie Collins, and Richard Newell. 2016. Distributed solar photovoltaic array location and extent dataset for remote sensing object identification. *Scientific Data* 3 (12 2016), 160106. https://doi.org/10.1038/sdata.2016.106

T. Chen, L. Zhu, C. Ding, R. Cao, Y. Wang, Z. Li, L. Sun, P. Mao, and Y. Zang. 2023. SAM Fails to Segment Anything? – SAM-Adapter: Adapting SAM in Underperformed Scenes: Camouflage, Shadow, Medical Image Segmentation, and More. (May 2023). arXiv:2304.09148 [cs]

Xin Hou, Biao Wang, Wanqi Hu, lei yin, Anbu Huang, and Haishan Wu. 2020. SolarNet: A Deep Learning Framework to Map Solar Plants In China From Satellite Imagery. In *ICLR 2020 Workshop on Tackling Climate Change with Machine Learning*. https://www.climatechange.ai/papers/iclr2020/6

S. Illarionova, D. Shadrin, I. Shukhratov, K. Evteeva, G. Popandopulo, N. Sotiriadi, I. Oseledets, and E. Burnaev. 2023. Benchmark for Building Segmentation on Up-Scaled Sentinel-2 Imagery. *Remote Sensing* 15, 9 (2023), 2347. https://doi.org/10.3390/rs15092347

M. Immitzer, F. Vuolo, and C. Atzberger. 2016. First Experience with Sentinel-2 data for crop and tree species classifications in Central Europe. *Remote Sensing* 8, 3 (2016), 166. https://doi.org/10.3390/rs8030166

H. Jiang, L. Yao, N. Lu, J. Qin, T. Liu, Y. Liu, and C. Zhou. 2021. Multi-resolution dataset for photovoltaic panel segmentation from satellite and aerial imagery. *Earth System Science Data* 13, 11 (2021), 5389–5401. https://doi.org/10.5194/essd-13-5389-2021

Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollár, and Ross Girshick. 2023. Segment Anything. arXiv:2304.02643

L. Kruitwagen, K. Story, J. Friedrich, L. Byers, S. Skillman, and C. Hepburn. 2021. A global inventory of photovoltaic solar energy generating units. *Nature* 598 (10 2021), 604–610. https://doi.org/10.1038/s41586-021-03957-7

Jordan Malof, Boning Li, Bohao Huang, Kyle Bradbury, and Artem Stretslov. 2019. Mapping solar array location, size, and capacity using deep learning and overhead imagery.

Kevin Mayer, Zhecheng Wang, Marie-Louise Arlt, Dirk Neumann, and Ram Rajagopal. 2020. DeepSolar for Germany: A deep learning framework for PV system mapping from aerial imagery. In *2020 International Conference on Smart Energy Systems and Technologies (SEST)*. 1–6. https://doi.org/10.1109/SEST48500.2020.9203258

A. Ortiz, D. Negandhi, S.R. Mysorekar, S.K. Nagaraju, J. Kiesecker, C. Robinson, P. Bhatia, A. Khurana, J. Wang, F. Oviedo, and J.L. Ferres. 2022. An Artificial Intelligence Dataset for Solar Energy Locations in India. *Scientific Data* 9, 1 (Aug 2022), 497. https://doi.org/10.1038/s41597-022-01499-9

Rafael Padilla, Sergio Netto, and Eduardo da Silva. 2020. A Survey on Performance Metrics for Object-Detection Algorithms. https://doi.org/10.1109/IWSSIP48289.2020

Simiao Ren, Francesco Luzi, Saad Lahrichi, Kaleb Kassaw, Leslie M. Collins, Kyle Bradbury, and Jordan M. Malof. 2023. Segment anything, from space? arXiv:2304.13000

J. Wu, W. Ji, Y. Liu, H. Fu, M. Xu, Y. Xu, and Y. Jin. 2023. Medical SAM Adapter: Adapting Segment Anything Model for Medical Image Segmentation. (2023). arXiv:2304.12620 [cs.CV]

Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. 2018a. DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States. *Joule* 2 (2018), 2605–2617.

Jiafan Yu, Zhecheng Wang, Arun Majumdar, and Ram Rajagopal. 2018b. DeepSolar: A Machine Learning Framework to Efficiently Construct a Solar Deployment Database in the United States. *Joule* 2, 12 (2018), 2605–2617. https://doi.org/10.1016/j.joule.2018.11.021

X. Zhu, D. Tuia, L. Mou, G.S. Xia, L. Zhang, F. Xu, and F. Fraundorfer. 2017. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience and Remote Sensing Magazine* 5, 4 (2017), 8–36. https://doi.org/10.1109/MGRS.2017.2762307

Li Zhuang, Zijun Zhang, and Long Wang. 2020. The automatic segmentation of residential solar panels based on satellite images: A cross learning driven U-Net method. *Appl. Soft Comput.* 92 (2020), 106283. https://api.semanticscholar.org/CorpusID:216382564