SOLAR PANEL SEGMENTATION: SELF-SUPERVISED LEARNING SOLUTIONS FOR IMPERFECT DATASETS

Sankarshanaa Sagaram*, **Krish Didwania***, **Laven Srivastava***, **Aditya Kasliwal***, **Pallavi Kailas*** & **Ujjwal Verma**†
*Manipal Institute of Technology, Manipal, †Manipal Institute of Technology Bengaluru
Manipal Academy of Higher Education, Manipal, Udupi, India
{sanky.sagaram,krish0674,lavensri,kasliwaladitya17}@gmail.com
kailaspallavi464@gmail.com, ujjwal.verma@manipal.edu

ABSTRACT

The increasing adoption of solar energy necessitates advanced methodologies for monitoring and maintenance to ensure optimal performance of solar panel installations. A critical component in this context is the accurate segmentation of solar panels from aerial or satellite imagery, which is essential for identifying operational issues and assessing efficiency. This paper addresses the significant challenges in panel segmentation, particularly the scarcity of annotated data and the labour-intensive nature of manual annotation for supervised learning. We explore and apply Self-Supervised Learning (SSL) to solve these challenges. We demonstrate that SSL significantly enhances model generalization under various conditions and reduces dependency on manually annotated data, paving the way for robust and adaptable solar panel segmentation solutions.

1 INTRODUCTION

The escalating role of solar energy in mitigating climate change has garnered increased research interest, propelled by technological advancements and heightened environmental awareness as portrayed in Zhuang et al. (2020). In this context, remote sensing has emerged as a pivotal tool for enhancing solar energy utilization, enabling the identification of regions with underutilized energy for targeted optimization (Rasmussen et al. (2021)). Nonetheless, segmenting solar panels with traditional supervised machine learning is challenging due to a lack of annotated datasets and the large amount of raw, unlabelled satellite data. Processing and accurately labelling this data requires extensive labour and is prone to inaccuracies, as shown by Fahy et al. (2022).

Our paper addresses segmentation challenges by leveraging extensive satellite imagery and using SSL techniques, bypassing the need for meticulously annotated data as in Chen et al. (2022). We demonstrate that SimCLR pretraining Chen et al. (2020) produces accurate segmentation masks despite label corruption. While pretraining incurs initial costs, it is marginal compared to fully annotating datasets. Our method reduces manual labour and overall costs, presenting a cost-effective and robust solution to challenges in applying machine learning to solar energy optimization Zhou (2017).

2 METHODOLOGY

We utilized the PV03 dataset (Jiang et al. (2021)), encompassing solar panel data from Jiangsu Province, China, and incorporated SimCLR pertaining as part of our experimental setup. Our experimentation involved various segmentation models, including encoder-decoder networks like U-Net(Ronneberger et al. (2015)) and FPN(Lin et al. (2017)), and PSPNet(Zhao et al. (2017)), which utilizes a pyramid pooling module. Additionally, we experimented with multiple backbones for each architecture, such as ResNets(He et al. (2015)), Mix Visual Transformer(Xie et al. (2021)), and VGG(Ronneberger et al. (2015)) with ImageNet pre-trained weights. The results of our best configurations are presented in 1. We employed Focal Loss (Lin et al. (2018)) for all configurations, which enhances learning by prioritising challenging misclassified examples. In all our experiments, we applied horizontal and vertical flips along with colour jittering at a probability of 0.5, and shifts in the HSV colour scale. Our analysis of these configurations led us to identify the most effective settings for our approach: a batch size of 8, focal loss parameters with an α value of 0.4 and a γ value of 2, a learning rate set at 3x10^{-5}, and the use of the Adam optimizer Kingma & Ba (2017) with β_1 and β_2 values of 0.9 and 0.99, respectively, during the fine-tuning process.

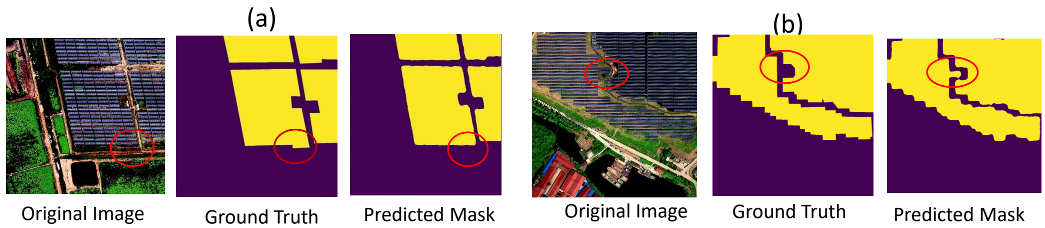

Figure 1: Instances demonstrating corrupted ground truths in the datasets, with contrast-enhanced RGB images for easier visualization, and how SSL adapts in the learning process.

3 Results

Initially, the entire training dataset was employed, followed by fine-tuning using varied subsets of this data before applying the model to the complete test set. Remarkably, diverse configurations, when applied to these subsets, consistently produced similar results during the fine-tuning stage. Notably, it achieved impressive performance even with limited annotated data as portrayed in Hendrycks et al. (2019), demonstrating that substantial annotation costs can be reduced by fine-tuning on a subset while still maintaining consistently high-quality results. Our pre-trained model was then fine-tuned on PSPNet with ResNet-34 encoder, following the methodology outlined in Jiang et al. (2021).

Table 1: PV03 Results with different configurations

	60%	**70%**	**80%**	**Full Dataset**
Unet [Jiang et al. (2021)]	-	-	-	0.858
PSPNet *	0.869	0.8841	0.869	0.8895
SimCLR +PSPNet *	0.8748	0.888	**0.8911**	0.890

Our study revealed two key findings: firstly, the superior performance of our predicted masks over the original, corrupted ground truths in the dataset, and secondly, penalties in terms of potentially lower IoU scores during the fine-tuning phase of experiments that were conducted with SSL pre-training. The predicted masks accurately segmented solar panels that were missed in the manual annotation process, as clearly demonstrated in Figure 1, where areas of interest are highlighted with red circles. This not only underscores the effectiveness of our methodology but also highlights the issue of label corruption in datasets. The lower IoU scores observed during SSL pre-training can be attributed to the generation of masks that more accurately represent real-world scenarios when compared to the provided ground truths, leading to unfair penalties during backpropagation, as seen in Figure1.

4 Conclusion

In our research, we have effectively applied SSL techniques to tackle the challenges of scarce and corrupted data annotations in solar panel segmentation. Our approach significantly improves model generalization, even under varying conditions, thereby addressing the limitations of manual annotation in supervised learning. The experiments conducted demonstrate that our SSL-based method consistently delivers strong performance, even with limited training data, highlighting its capability to mitigate the issues of label scarcity and corrupted labels. Furthermore, the enhanced robustness achieved in this domain is a testament to the versatility of our methodology.

URM STATEMENT

The authors acknowledge that all key authors of this work meet the URM criteria of ICLR 2024 Tiny Papers Track.

REFERENCES

De-Yue Chen, Ling Peng, Wen-Yue Zhang, Yin-Da Wang, and Li-Na Yang. Research on self-supervised building information extraction with high-resolution remote sensing images for photovoltaic potential evaluation. *Remote Sensing*, 14(21), 2022. ISSN 2072-4292. doi: 10.3390/rs14215350. URL https://www.mdpi.com/2072-4292/14/21/5350.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In Hal Daumé III and Aarti Singh (eds.), *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 1597–1607. PMLR, 13–18 Jul 2020. URL https://proceedings.mlr.press/v119/chen20j.html.

Conor Fahy, Shengxiang Yang, and Mario Gongora. Scarcity of labels in non-stationary data streams: A survey. *ACM Comput. Surv.*, 55(2), jan 2022. ISSN 0360-0300. doi: 10.1145/3494832. URL https://doi.org/10.1145/3494832.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.

Dan Hendrycks, Mantas Mazeika, Saurav Kadavath, and Dawn Song. Using self-supervised learning can improve model robustness and uncertainty. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper_files/paper/2019/file/a2b15837edac15df90721968986f7f8e-Paper.pdf.

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

Maxim Khomiakov, Julius Holbech Radzikowski, Carl Anton Schmidt, Mathias Bonde Sørensen, Mads Andersen, Michael Riis Andersen, and Jes Frellsen. Solardk: A high-resolution urban solar panel image classification and localization dataset. *arXiv preprint arXiv:2212.01260*, 2022.

Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, and Serge Belongie. Feature pyramid networks for object detection, 2017.

Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection, 2018.

Jesper Rasmussen, Saiful Azim, Søren Kjærgaard Boldsen, Thomas Nitschke, Signe M Jensen, Jon Nielsen, and Svend Christensen. The challenge of reproducing remote sensing data from satellites and unmanned aerial vehicles (uavs) in the context of management zones and precision agriculture. *Precision Agriculture*, 22:834–851, 2021.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo. Segformer: Simple and efficient design for semantic segmentation with transformers, 2021.

Zhilu Zhang and Mert R. Sabuncu. Generalized cross entropy loss for training deep neural networks with noisy labels, 2018.

Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia. Pyramid scene parsing network, 2017.

Zhi-Hua Zhou. A brief introduction to weakly supervised learning. *National Science Review*, 5(1): 44–53, 08 2017. ISSN 2095-5138. doi: 10.1093/nsr/nwx106. URL `https://doi.org/10.1093/nsr/nwx106`.

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

A LOSS FUNCTION

We employ the use of the Focal Loss Lin et al. (2018) for the fine-tuning stages of our approach, introducing a modulating term to the cross-entropy Zhang & Sabuncu (2018) loss, prioritise learning on challenging misclassified examples. The adjustable hyperparameters, α and γ enable us to control the focusing effect and the rate at which the loss decreases for well-classified sample . The tuning parameter gamma governs the rate at which simple examples are down-weighted alongside alpha, which plays a pivotal role in addressing class imbalance by weighting the loss of each class. This dynamic scaling of the cross entropy loss involves the scaling factor diminishing to zero as confidence in the correct class rise . In image segmentation tasks, the Focal loss assumes a pivotal role in handling class imbalance and accentuating challenging samples, resulting in notable improvements in segmentation performance. The Focal loss formula is given as follows:

$$\text{Focal Loss} = -\frac{1}{n}\sum_{i=1}^{n}\left(\alpha(1-\hat{y}_i)^{\gamma} y_i \log(\hat{y}_i) + (1-\alpha)\hat{y}_i^{\gamma}(1-y_i)\log(1-\hat{y}_i)\right) \tag{1}$$

In the formula, y_i represents the ground truth value (0 or 1) for the i-th sample, and $\hat{y}_i$ represents the corresponding predicted value from the mode . The summation runs over all n samples in the dataset.

B PRETRAINING AS A METHOD TO GENERALIZE ACROSS DATASETS

B.1 SOLARDK DATASET

We have also used another dataset, SolarDK Khomiakov et al. (2022), which encompasses labelled instances from two urban municipalities in the Greater Copenhagen Region and the Danish Building Registry (BBR), to evaluate and test the generalizability of our approach across datasets. As demonstrated in the experimentation section, our approach effectively addresses the robustness and adaptability of SSL in navigating the challenges inherent in solar panel segmentation and validation.

B.2 GENERALIZABILITY

Notably, as demonstrated in Table 2, our findings indicate that conducting pre-training and finetuning on two distinct datasets yields superior results compared to performing both operations on the same dataset. This underscores the resilience and impressive generalization capabilities of self-supervised algorithms within this domain.

As demonstrated in the experimentation section, our approach effectively addresses the robustness and adaptability of SSL in navigating the challenges inherent in solar panel segmentation and validation. We further enhanced our methodology by implementing cross-dataset validation on SimCLR, wherein the model is pre-trained on solarDK and fine-tuned on PV03 and vice versa.

Table 2: For cross-validation, a U-net model was used alongside the ResNet-34 backbone for SolarDk; PSPNet with the ResNet-34 fine-tuned the PV03 dataset.

Pre-training	**Fine-tuning**	**Max IoU**
SolarDK	SolarDK	0.649
PV03	**SolarDK**	**0.6971**
PV03	PV03	0.8829
SolarDK	**PV03**	**0.8928**

This method was employed to evaluate the robustness of pre-training in this domain, aiming to highlight the potential of extracting solar panel information and features across similar datasets. As demonstrated in Table 2, our findings indicate that conducting pre-training and fine-tuning on two distinct datasets yields superior results compared to performing both operations on the same dataset. This underscores the resilience and improved generalisation capabilities of self-supervised algorithms within this domain.